

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: February 12, 2003, 11:39:30 ; Search time 32 Seconds  
(without alignments)  
33.313 Million cell updates/sec

Title: US-09-660-302c-1  
Perfect score: 8  
Sequence: 1 XXXXXXXX 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq-101002 : \*

- 1: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1980.DAT:\*
- 2: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1981.DAT:\*
- 3: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1982.DAT:\*
- 4: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1983.DAT:\*
- 5: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1984.DAT:\*
- 6: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1985.DAT:\*
- 7: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1986.DAT:\*
- 8: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1987.DAT:\*
- 9: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1988.DAT:\*
- 10: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1989.DAT:\*
- 11: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1990.DAT:\*
- 12: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1991.DAT:\*
- 13: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1992.DAT:\*
- 14: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1993.DAT:\*
- 15: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1994.DAT:\*
- 16: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1995.DAT:\*
- 17: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1996.DAT:\*
- 18: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1997.DAT:\*
- 19: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1998.DAT:\*
- 20: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA1999.DAT:\*
- 21: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA2000.DAT:\*
- 22: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA2001.DAT:\*
- 23: /SIDS2/gcdatadata/geneseq/geneseqp-embl/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	0	0 0	1 20	AAY46552	Immunogenic peptide
2	0	0 0	1 22	ABB6670	Human SNP related
3	0	0 0	1 22	ABB6609	Drosophila melanog
4	0	0 0	1 22	ABB6610	Drosophila melanog
5	0	0 0	1 22	ABG0241	Novel human diagno
6	0	0 0	1 22	AM97043	Human peptide #118
7	0	0 0	1 22	AAM9734	Human peptide #110
8	0	0 0	1 22	AAM9774	Human peptide #124
9	0	0 0	1 22	AAM98354	Human peptide #162
10	0	0 0	1 22	AAM98447	Human peptide #172

Synthetic.

OS Homo sapiens.

XX WO9945954-A1.

XX 16-SEP-1999.

XX 01-DEC-1999 (first entry)

XX AAY46552

ID AAY46652 standard; Peptide; 1 AA.

XX

AC AAY46552;

XX

DT 13-MAR-1998;

XX 98WO-US05039.

PR 13-MAR-1998;

XX 98WO-US05039.

PA (EPIIM ) EPIIMMUNE INC

XX

P1 Sette, A., Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;

XX DR WPI: 1999-551214/46.

New immunogenic peptides with HLA binding motif, useful in treatment and diagnosis of cancers and viral diseases -

Claim 1; Page 80; 150pp; English.

AAV45390 to AAV48214 represent specifically claimed immunogenic peptides having a human major histocompatibility complex (MHC) Class I (also known as human leukocyte antigen (HLA)) binding motif. The immunogenic peptides can bind to a specific HLA allele (i.e., HLA-A subtypes HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T-cell response against the antigen from which the peptide is derived. Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are normally induced by an antigen in the form of a peptide fragment bound to a HLA molecule, rather than the intact foreign antigen itself, and are particularly important in tumour rejection and in fighting viral infections. The peptides are therefore useful therapeutically to treat or prevent viral infections and cancers in mammals (especially humans e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma). They can be administered as vaccines to elicit an immune response in individuals susceptible or otherwise at risk of viral infection or cancer, or used to treat chronic or acute conditions. They are also useful diagnostically, and can be used to induce a cytotoxic T-cell response, by contacting a cytotoxic T cell with the peptide e.g. to produce CTLs ex vivo for infusion back into a patient. The polynucleotides encoding the immunogenic peptides are also useful therapeutically and for immunisation as above.

XX	Claim 1; Page 666; 674pp; English.
PS	ABL00010 to ABL01104 represent human nucleic acid oligonucleotides comprising one or more single nucleotide polymorphisms (SNPs). ABB56531 to ABB56903 represent human peptides encoded by some of the SNP Oligonucleotides. The sequences from the present invention can have immunosuppressive, cytostatic, antiinflammatory, neuroprotective and antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides and antibodies from the present invention can be used for treating a subject suffering from, at risk for, or suspected of suffering from a pathology ascribed to the presence of a sequence polymorphism. The pathology may be autoimmune diseases, inflammation, cancer, diseases of the nervous system, and infection by pathogenic microorganisms. The SNPs are also useful for determining which forms of a characterised polymorphism are present in individuals. The antibodies may be used in the detection, quantitation and/or cellular or tissue localisation of a polymorphic protein (e.g., for use in measuring levels of the polymorphic protein within appropriate physiological samples).
XX	SQ Sequence 1 AA;
XX	Query Match Score 0; Best Local Similarity 0.0%; Length 1; Matches 0; Conservative 0; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0
QY	1 X 1
Db	1 S 1
RESULT 3	
XX	ABB66809 standard; Protein; 1 AA.
XX	AC ABB66809;
XX	DT 26-MAR-2002 (first entry)
XX	DE Drosophila melanogaster polypeptide SEQ ID NO 27219.
XX	KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.
XX	OS Drosophila melanogaster.
XX	PN WO200171042-A2.
XX	PD 27-SEP-2001.
XX	PF 23-MAR-2001; 2001WO-US09231.
XX	PR 23-MAR-2000; 2000US-191637P. 11-JUL-2000; 2000US-0614150.
XX	PA (PEKE ) PE CORP NY.
PI	venter JC, Adams M, Li PWD, Myers EW;
XX	DR 2001-65860/75. N-PSDB; ABL10912.
XX	PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
XX	PS Disclosure; SEQ ID NO 27219; 21pp + Sequence Listing; English.
XX	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention includes

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABB7737-ABB7072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at [ftp://wipo.int/pub/published\\_pct\\_sequences](ftp://wipo.int/pub/published_pct_sequences).

XX Sequence 1 AA;

SQ

Query Match 0.0%; Score 0; DB 22; Length 1;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 x 1

Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy 1 x 1  
 Db 1 H 1

Qy





XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic; neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer; amyloid protein; angiopoietin; apoptosis related protein; cadherin; cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor; complement related protein; cytochrome; kinesin; cytokine; interferon; interleukin; G-protein coupled receptor; thioesterase; inflammation; multifactorial disease; autoimmune disease; infection; nervous system disease.

XX Homo sapiens.

OS WO200140521-A2.

XX 05-JUL-2001.

XX 28-DEC-2000; 2000WO-US35498.

XX 28-DEC-1999; 99US-0173419.

XX 27-DEC-2000; 2000US-0173419.

PA (CURA-) CURAGEN CORP.

XX Shinkets RA, Leach M;

DR WPI; 2001-465210/50.

XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases, oncogenes and histones, useful for diagnosing and treating, e.g. cancer, autoimmune diseases and infections.

XX Disclosure: Page 4045; 4143pp; English.

The present invention relates to oligonucleotides (see AAI26793-AAI34659) encoding polymorphic variants of proteins related to amylases, amyloid proteins, angiopoietin, apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes, histones, kinases, colony stimulating factors, complement related proteins, cytochromes, kinesins, cytokines, interleukins, G-protein coupled receptors and thioesterases. The present sequence is a peptide encoded by one such oligonucleotide, the oligonucleotides and the peptides encoded by them may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate expression of the proteins listed above. Disorders that may be prevented, diagnosed and/or treated include multifactorial diseases with a genetic component, such as autoimmune diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus and Grave's disease), inflammation, cancer (e.g. cancers of the bladder, brain, breast, colon and kidney, leukaemia), diseases of the nervous system and an infection of pathogenic organisms.

Sequence 1 AA;

Query	Match	Score	Length	Pred.	No.	DB	22;	Score	0;	DB	22;	Length	1;											
QY	1 X 1	0.0%	1;	Mismatches	0;	Indels	0;	Gaps	0;	Matches	0;	Conservative	0;	Best Local Similarity	0.0%;	Pred.	No.	0;	Mismatches	1;	Indels	0;	Gaps	0;
Db	1 L 1																							

Query Match Score 0; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Sequence 1 AA;

Query Match Score 0.0%; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Sequence 1 AA;

Query Match Score 0.0%; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Sequence 1 AA;

Query Match Score 0.0%; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Sequence 1 AA;

Query Match Score 0.0%; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Sequence 1 AA;

Query Match Score 0.0%; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Sequence 1 AA;

Query Match Score 0.0%; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Sequence 1 AA;

Query Match Score 0.0%; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Sequence 1 AA;

Query Match Score 0.0%; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Sequence 1 AA;

Query Match Score 0.0%; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Sequence 1 AA;

Query Match Score 0.0%; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Sequence 1 AA;

Query Match Score 0.0%; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Sequence 1 AA;

Query Match Score 0.0%; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Sequence 1 AA;

Query Match Score 0.0%; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Sequence 1 AA;

Query Match Score 0.0%; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Sequence 1 AA;

Query Match Score 0.0%; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Sequence 1 AA;

Query Match Score 0.0%; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Sequence 1 AA;

Query Match Score 0.0%; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Sequence 1 AA;

Query Match Score 0.0%; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Sequence 1 AA;

KW quantitation; restorative therapy; polymorphic.

XX Homo sapiens.

OS WO200140521-A2.

XX 07-JUN-2001.

XX 30-NOV-2000; 2000WO-US32758.

XX 30-NOV-1999; 99US-016138.

PR 29-NOV-2000; 2000US-0726173.

XX (CURA-) CURAGEN CORP.

XX Shinikets RA, Leach M;

PS Claim 29; Page 2619; 2653pp; English.

XX AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide sequences (SNPs), which contain single nucleotide polymorphisms (SNPs).

CC AAM53114 to AAM53329 represent peptides related to human polymorphic polynucleotide sequences. The sequences can be used in gene and protein therapy, and in vaccine production. (1) and the polypeptides encoded by them may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate expression of polymorphic polypeptides. For example, (1) may be used to treat disorders by rectifying mutations or deletions in a patient's genome that affect the activity of polypeptides by expressing inactive proteins or to supplement the patient's own production of polypeptide. Additionally, (1) and its complementary sequences may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acids in samples, and therefore which patients may be in need of restorative therapy. The polypeptides encoded by (1) may be used as antigens in the production of antibodies specific for polymorphic polypeptides. The antibodies may also be used to down regulate expression and activity. The antibodies may also be used as diagnostic agents for detecting the presence of polymorphic polypeptides in samples.

CC Sequence 1 AA;

CC Query Match Score 0.0%; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Score 0;

CC DB 1 H 1

CC RESULT 12

CC AAM53219 standard; Peptide; 1 AA.

CC ID AAM53219

CC AC AAM53219;

CC DT 09-NOV-2001 (first entry)

XX Human nonconservative amino acid changing SNP related peptide SEQ:6914.

XX DE Human nonconservative amino acid changing SNP related peptide SEQ:6914.

XX KW Human; single nucleotide polymorphism; SNP; genome; gene therapy; protein therapy; vaccine; probe; diagnostic assay; detection; quantitation; restorative therapy; polymorphic.

XX Homo sapiens.

PN WO200140521-A2.

PD 07-JUN-2001.  
 XX 30 -NOV-2000; 2000WO-US32258.  
 PF XX  
 PR 30 -NOV-1999; 99US-0168138.  
 PR 29 -NOV-2000; 2000US-0726173.  
 XX PA (CURA-) CURAGEN CORP.  
 XX PI Shimkets RA, Leach M;  
 XX DR; 2001-356160/37.  
 PT Polymorphic nucleic acid sequences useful in genetic testing and therapy -  
 XX PS Claim 29; Page 2619; 2653pp; English.  
 XX AA173060 to AA179867 represent isolated human polymorphic polynucleotide sequences (I), which contain single nucleotide polymorphisms (SNPs).  
 CC AA173060 to AA179867 represent isolated human polymorphic polynucleotide sequences (I), which contain single nucleotide polymorphisms (SNPs).  
 CC AAM53329 represent peptides related to human polymorphic polynucleotide sequences. The sequences can be used in gene and protein therapy, and in vaccine production. (I) and the polypeptides encoded by them may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate expression of polymorphic polypeptides. For example, (I) may be used to treat disorders by rectifying mutations or deletions in a patient's genome that affect the activity of polypeptides by expressing inactive proteins or to supplement the patients own production of polypeptide. Additionally, (I) and its complementary sequences may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acids in samples, and therefore which patients may be in need of restorative therapy. The polypeptides encoded by (I) may be used as antigens in the production of antibodies specific for polymorphic polypeptides. The antibodies may also be used to down regulate expression and activity. The antibodies may also be used as diagnostic agents for detecting the presence of polymorphic polypeptides in samples.

XX SQ Sequence 1 AA;

Query Match Score 0; DB 22; Length 1;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
 Db 1 L 1

RESULT 14  
 ID AAM53391 standard; Peptide: 1 AA.  
 XX ID AAM53391;  
 AC AAM533291;  
 XX DT 09-NOV-2001 (first entry)  
 XX DT 09-NOV-2001 (first entry)

RESULT 13  
 ID AAM53390 standard; Peptide: 1 AA.  
 XX AC AAM533290;  
 XX DT 09-NOV-2001 (first entry)  
 XX DE Human nonconservative amino acid changing SNP related peptide SEQ:6985.  
 XX KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;  
 KW protein therapy; vaccine; probe; diagnostic assay; detection;  
 KW quantitation; restorative therapy; polymorphic.  
 XX OS Homo sapiens.  
 XX PF 30 -NOV-2000; 2000WO-US32258.  
 XX PN WO200140521-A2.  
 XX PR 30 -NOV-1999; 99US-0168138.  
 XX PD 29 -NOV-2000; 2000US-0726173.  
 XX PA (CURA-) CURAGEN CORP.  
 XX PI Shimkets RA, Leach M;  
 XX DR; 2001-356160/37.

```

Sequence 1 AA;
Query Match 0.0%; Score 0; DB 22; Length 1;
Best Local Similarity 0.0%; Pred. NO. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 x 1

Qy 1 x 1
Db 1 v 1

Search completed: February 12, 2003, 11:44:37
Job time : 33 secs

```

RESULTS 15  
AM53328 standard; Peptide: 1 AA.  
AM53328;

09-NOV-2001 (first entry)  
Human nonconservative amino acid changing SNP related peptide SEQ:7023.  
Human; single nucleotide polymorphism; SNP; genome; gene therapy;  
protein therapy; vaccine; probe; diagnostic assay; detection;  
quantitation; restorative therapy; polymorphic

Homo sapiens.

WO200140521-A2.  
07 -JUN-2001.  
30 -NOV-2000; 2000WO-US32758.  
30 -NOV-1999; 99US-0168138.  
29 -NOV-2000; 2000US-0726173.

(CURA-) CURAGEN CORP.  
Shimkets RA, Leach M;  
WPI; 2001-356160/37.  
Polymorphic nucleic acid sequences, useful in genetic testing and  
therapy -  
Claim 29; Page 2652; 2653pp; English.

GenCore version 5.1.3  
 Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: February 12, 2003, 11:39:59 ; Search time 7.46667 Seconds  
 (without alignments)

Title: US-09-660-302c-1  
 Perfect score: 8  
 Sequence: 1 XXXXXXXX 8

Scoring table: Blosum62  
 Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : SwissProt\_40: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	0	0	3	1	GRWM_HUMAN	P01157 homo sapien
2	0	0	3	1	LUXE_VIBFI	P24272 vibrio fiscus
3	0	0	3	1	THYL_PIGF	P01151 suis scrofa
4	0	0	4	1	ACH1_ACHFEU	P35904 achatinina fumosa
5	0	0	4	1	DCML_PSECH	P19916 pseudomonas
6	0	0	4	1	DCMS_PSECH	P19918 pseudomonas
7	0	0	4	1	EOSI_HUMAN	P02731 homo sapiens
8	0	0	4	1	FAR3_HIRME	P42562 hirudo medicinalis
9	0	0	4	1	FAR4_HIRME	P42563 hirudo medicinalis
10	0	0	4	1	FFKA_ANTEL	P58705 anthropoileurus
11	0	0	4	1	FLRF_HIRME	P42561 hirudo medicinalis
12	0	0	4	1	FLRN_ANTEL	P58707 anthropoleura
13	0	0	4	1	FMRF_MACN1	P01162 macrocallis
14	0	0	4	1	FYRI_ANTEL	P58706 anthropoleura
15	0	0	4	1	OCP1_OCTMI	P58648 octopus minor
16	0	0	4	1	OCP3_OCTMI	P58649 octopus minor
17	0	0	4	1	RM01_YEAST	P36515 sachcharomyces
18	0	0	4	1	TUFT_HUMAN	P01838 homo sapiens
19	0	0	5	1	ALL4_CARMA	P81817 carcinoma
20	0	0	5	1	BIOA_CITFR	P13071 citrobacter
21	0	0	5	1	BIOB_CITFR	P12997 citrobacter
22	0	0	5	1	BPP7_BOTIN	P30425 bothrops inaequinoctialis
23	0	0	5	1	E103_LITRU	P82039 litoria rubra
24	0	0	5	1	E104_LITRU	P82100 litoria rubra
25	0	0	5	1	FARP_ARTTR	P41853 artiopositha
26	0	0	5	1	PAP2_PARMA	P81864 pardalotinus
27	0	0	5	1	PRCT_PERM	P01373 periplaneta americana
28	0	0	5	1	PSK_DAICA	P58261 daucus carota
29	0	0	5	1	RE11_LITRU	P82070 litoria rubra
30	0	0	5	1	RE21_LITRU	P82071 litoria rubra
31	0	0	5	1	RE31_LITRU	P82072 litoria rubra
32	0	0	5	1	RE32_LITRU	P82073 litoria rubra
33	0	0	5	1	SUGA_ACHDO	P19991 acheta domesticus

## ALIGNMENTS

RESULT 1	GRWM_HUMAN	ID	GRWM_HUMAN	STANDARD;	PRT;	3 AA.
		AC	P01157;			
		DT	21-JUL-1986 (Rel. 01, Created)			
		DT	21-JUL-1986 (Rel. 01, Last sequence update)			
		DT	21-JUL-1986 (Rel. 01, Last annotation update)			
		DE	Growth-modulating Peptide.			
		OS	Homo sapiens (Human).			
		OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
		OX	NCBI_TaxID=9606;			
		RN	[1..]			
		RP	SEQUENCE.			
		RX	Medline=77162369; PubMed=8583556;			
		RA	Schliebsinger D.H.; Pickart L.; Thaler M.M.;			
		RL	"Growth-modulating serum tripeptide is glycyl-histidyl-lysine."			
		RT	Experimental 33:324-325 (1977).			
		CC	-- MISCELLANEOUS: THIS SERUM TRIPETIDE HAS BEEN FOUND TO STIMULATE GROWTH OF SOME CELL TYPES AND TO INHIBIT OTHER TYPES IN VITRO.			
		CC	PIR; A01421; GKHU.			
		DR	SEQUENCE 3 AA; 340 MW; 6331E810000000000 CRC64;			
		Qy	1 X 1			
		Db	1 G 1			
		DR	1 G 1			
		LUXE_VIBFI	Score 0; DB 1; Length 3;			
		AC	Best Local Similarity 0.0%; Pred. No. 0;			
		AC	Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
		DT	Query Match			
		DT	Best Local Similarity			
		DT	Matches			
		DT	Conservative			
		DT	Mismatches			
		DT	Indels			
		DT	Gaps			
		RT	SEQUENCE FROM N.A.			
		RT	Medline=91072226; PubMed=2254256;			
		RA	Swartzman B., Kapoor S., Graham E.A.; Meighen E.A.;			
		RA	"A new Vibrio fischeri lux gene precedes a bidirectional termination site for the lux operon."			
		RT	J. Bacteriol. 172:6797-6802 (1990).			
		RL	-- FUNCTION: ACYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID.			
		CC	IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE FOR CONVERTING TETRADECANOIC ACID TO THE ALDHYDE WHICH SERVES AS SUBSTRATE IN THE LUCIFERASE-CATALYZED REACTION.			

-!- CATALYTIC ACTIVITY: ATP + an acid + protein = AMP + diphosphate +  
 an acyl-protein thioester.  
 -!- PATHWAY: Bioluminescent fatty acid reduction system; second step.  
  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on i  
 use by non-profit institutions as long as its content is in no w  
 modified and this statement is not removed. Usage by and for commer  
 enties requires a license agreement (See http://www.isb-sib.ch/announc  
 or send an email to license@isb-sib.ch).  
  
 EMBL; M62812; ; NOT\_ANNOTATED\_CDS.  
 Luminescence; Ligase; ;  
 NON\_TER 1 1  
 SEQUENCE 3 AA: 374 MW: 66A3303000000000 CRC64;  
  
 Query Match 0.0%; Score 0; DB 1; Length 3;  
 Besty Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Sus scrofa (Pig);  
 Os avis (Sheep);  
 Bombina orientalis (Oriental fire-bellied toad), and  
 Notophthalmus viridescens (Eastern newt) (*Triturus viridescens*);  
 Nair R.M.G., Barrett J.F., Bowers C.Y., Schally A.V.;  
 Mammalia; Eutheria; Chordata; Craniota; Vertebrata; Euteleostomi;  
 NCBI\_TAXID=9823, 9940, 8346, 8316;  
 [1]  
 SEQUENCE.  
 SPECIES=Pig; TISSUE=Hypothalamus;  
 MEDLINE=70136150; PubMed=4984938;  
 Bojer J., Enzmann F., Folkers K., Bowers C.Y., Schally A.V.;  
 "Structure of porcine thyrotropin releasing hormone";  
 Biochemistry 9:1103-1106(1970).  
 [2]  
 SYNTHESIS  
 SPECIES=Pig;  
 MEDLINE=70139904; PubMed=4982117;  
 Bojer J., Enzmann F., Folkers K., Bowers C.Y., Schally A.V.;  
 "The identity of chemical and hormonal properties of the thyrotropin  
 releasing hormone and pyroglutamyl-histidyl-proline amide";  
 Biochem. Biophys. Res. Commun. 37:705-710(1969).  
 [3]  
 SEQUENCE.  
 SPECIES=Sheep; TISSUE=Hypothalamus;  
 Desiderio D.M. Jr., Burgus R., Dunn T.F., Vale W., Guillemin R.,  
 Ward D.N.;  
 "The elucidation of the primary structure of the hypothalamic thyroid  
 stimulating hormone releasing factor of ovine origin by means of mass  
 spectrometry";  
 J. Org. Mass Spectrom. 5:221-228(1971).  
 [4]  
 SYNTHESIS  
 SPECIES=Sheep;  
 MEDLINE=701386; PubMed=4985794;  
 Burgus R., Dunn T.F., Desiderio D.M., Ward D.N., Vale W.,  
 Guillemin R.;  
 "Characterization of ovine hypothalamic hypophysiotropic  
 TSH-releasing factor";  
 J. Org. Mass Spectrom. 5:221-228(1971).

Nature 226: 321-325 (1970).

RL Nature 226:321-325(1970).  
 RN [5]  
 RP SEQUENCE.  
 RC SPECIES=B orientalis; TISSUE=Skin;  
 RM MEDLINE=76138399; PubMed=815011;  
 RX

between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce> or send an email to licensed@isb-sib.ch).

```

C C
- - - - -
C E M B L ; M 6 2 8 1 2 ; - N O T _ A N N O T A T E D _ C D S .
W R
L u m i n e s c e n c e ; L i g a s e .
T T
N O N _ T E R 1 1
S E Q U E N C E 3 A A ; 3 7 4 M W ; G A A 3 3 0 3 0 0 0 0 0 0 0 0 0 0 0 0 C R C 6 4 ;
Q Q

Query Match 0 . 0 % ; Score 0 ; DB 1 ; Length 3 ;
Best Local Similarity 0 . 0 % ; Pred . No . 0 ;
Matches 0 ; Conservative 0 ; Mismatches 1 ; Indels 0 ; Caps

```

NCBI\_TAXID=9823; 9940; 8346; 8316;  
 [1]  
 SEQUENCE  
 SPECIES=Pig; TISSUE=Hypothalamus;  
 MEDLINE=70136150; PubMed=4984938;  
 Nair R.M.G., Barrett J.F., Bowers C.Y., Schally A.V.;  
 "Structure of porcine oxytocin releasing hormone.",  
 Biochemistry 9:1103-1110(1970).

[2] SPECIES=Pig; MEDLINE=70039904; PubMed=4982117; Bowers C.Y., Follers J., Boier J., Enzmann D.N.; "The identity of chemical and hormonal properties of the thyrotropin releasing hormone and pyroglutamyl-histidyl-proline amide."; Biochem. Biophys. Res. Commun. 37:705-710(1969).

[3] SPECIES=Sheep; TISSUE=Hypothalamus; Desiderio D.M. Jr., Burgus R., Dunn T.F., Vale W., Guillemin R., Ward D.N.; "The elucidation of the primary structure of the hypothalamic thyroid stimulating hormone releasing factor of ovine origin by means of mass sequencing."; J. Biol. Chem. 250:10353-10358(1975).

specify; Org. Mass Spectrom. 5:221-228(1971).  
 N [4]  
 P SYNTHESIS.  
 C SPECIES=Sheep;  
 X MEDLINE=70163386; PubMed=4985794;  
 A Burgos R., Dunn T.F., Desiderio D.M., Ward D.N., Vale W.,  
 A Guillemin R.;  
 A "Characterization of ovine hypothalamic hypophysiotropic  
 T TSH-releasing factor.";

RJ Nature 226:321-325(1970).

[5]

RN RP SEQUENCE.

RC SPECIES=B.orientalis; TISSUE=Skin;

RX MEDLINE=76138399; PubMed=815011;

RA Yasuhara T.; Nakajima T.;

RT Letter: Occurrence of Pyr-His-Pro-NH<sub>2</sub> in the frog skin.";

RJ Chem. Pharm. Bull. 23:3301-3303(1975).

[6]

RN RP SEQUENCE.

RC SPECIES=N. viridescens;

RX MEDLINE=75035605; PubMed=214528;

RA Grimm-Joergensen Y.; McKelvy J.F. "Biosynthesis of thyrotropin releasing factor by newt (*Triturus viridescens*) brain in vitro". Isolation and characterization of thyrotropin releasing factor.".

RJ J. Neurochem. 23:471-478(1974).

-1- FUNCTION: TRH FUNCTIONS AS A REGULATOR OF THE BIOSYNTHESIS OF T IN THE ANTERIOR PITUITARY GLAND AND AS A NEUROTRANSMITTER/NEUROMODULATOR IN THE CENTRAL AND PERIPHERAL NERVOUS SYSTEMS.

CC CC

RESULT 4						
ACHL_ACIFU	STANDARD:	PRT;	4 AA.			
ID ACHL_ACHFU						
AC P35404;						
DT 01-JUN-1994	(Rel. 29, Created)					
DT 01-JUN-1994	(Rel. 29, Last sequence update)					
DT 15-JUL-1998	(Rel. 36, last annotation update)					
DE Achatina-T						
OS Achatina fulica (Giant African snail).						
OC Achatinaceae; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora; Achatinidae; Achatina.						

OX NCBI\_TaxID=6330;  
RN RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.  
RC SPEAIN=Ferussac; TISSUE=Ganglion;  
MEDLINE=89273551; PubMed=2597281;  
RX RA Kanatani Y., Minakata H., Kenny P.T.M., Iwashita T., Watanabe K.,  
RA Funase K., Sun X.P., Yongsiri A., Kim K.H., Novales-Li P.,  
RA Novales E.T., Kanapi C.G., Takeuchi H., Nomoto K.;  
RA "Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina fulica Ferussac containing a D-amino acid residue";  
RT RL Biochem. Biophys. Res. Commun. 160:1015-1020(1989).  
RN [1]  
RP CHARACTERIZATION,  
RC STRAIN=Ferussac; TISSUE=Heart atrium;  
RX MEDLINE=91264856; PubMed=1675568;  
RA Fujimoto K., Kubota I., Yasuda-Kamatani Y., Minakata H., Nomoto K.,  
RA Yoshida M., Harada A., Muneoka Y., Kobayashi M.;  
RT "Purification of achatin-I from the atria of the African giant snail Achatina fulica, and its possible function";  
RT RL Biochem. Biophys. Res. Commun. 177:847-853(1991).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY.  
RX MEDLINE=93014529; PubMed=1399265;  
RA Ishida T., Inoue Y., Doi M., Inoue M., Yasuda-Kamatani Y., Minakata H.,

RA	Iwashita T., Nomoto K.; "Crytall structure and molecular conformation of achatin-I (H-Gly-D-Phe-Ala-Asp-OH), an endogenous neuropeptide containing a D-amino acid residue."; J. Pept. Protein Res. 39:258-264 (1992).	dehydrogenase subunit S) (CO-DH S) (Fragment).
RT	-!- FUNCTION: NEUROEXCITATORY PEPTIDE; INCREASES THE IMPULSE FREQUENCY AND PRODUCES A SPIKE BROADENING OF THE IDENTIFIED HEART EXCITATORY NEURON (PON); ALSO ENHANCES THE AMPLITUDE AND FREQUENCY OF THE HEART BEAT HAS ALSO AN EFFECT ON SEVERAL OTHER MUSCLES.	PubMED=2818128;
DR	PIR; A32480; A32480.	RX MEDLINE=90055678;
KW	KW Hormone; D-amino acid.	RA Kraut M., Hugendieck I., Herwig S., Meyer O.; OC Homology and distribution of CO dehydrogenase structural genes in carboxydrophic bacteria."
MOD_RES	FT 2	RT Arch. Microbiol. 152:335-341(1989);
SEQUENCE	SQ 4 AA; 408 MW; D-PHENYLALANINE. 6AADD9C81000000 CRC64;	RL -!- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon dioxide.
Qy	Query Match 0.0%; Score 0; DB 1; Length 4;	CC -!- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced acceptor.
Db	Db Best Local Similarity 0.0%; Pred. No. 0; Indels 0; Gaps 0; Matches 0; Conservative 0; Mismatches 1; Indels 1; Gaps 0;	CC -!- COFACTOR: BINDS TWO 2FE-2S CLUSTERS.
	Qy 1 X 1	CC -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND SMALL.
	Db 3 A 3	DR PIR: PL0146; PL0146.
		KW Oxidoreductase; Iron-sulfur.
		FT NON-TER 4 4 MW; 6DD33DD6F0000000 CRC64;
		SQ SEQUENCE 4 AA; 420 MW;
		Query Match 0.0%; Score 0; DB 1; Length 4;
		Best Local Similarity 0.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
		Matches 0; Conservative 0; Mismatches 1; Indels 1; Gaps 0;
		QY 1 X 1
		Db 2 A 2
		RESULT 7 EOSI_HUMAN STANDARD; PRT; 4 AA.
		ID EOSI_HUMAN STANDARD; PRT; 4 AA.
		AC P02731;
		DT 21-JUL-1986 (Rel. 01, Created)
		DT 21-JUL-1986 (Rel. 01, Last sequence update)
		DT 21-JUL-1986 (Rel. 01, Last annotation update)
		DE Eosinophilic Peptides.
		OS Homo sapiens (Human).
		OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
		OX NCBI_TAXID=9606;
		RN
		SEQUENCE.
		RX MEDLINE=90055678; PubMED=2818128;
		RA Kraut M., Hugendieck I., Herwig S., Meyer O.; OC Homology and distribution of CO dehydrogenase structural genes in carboxydrophic bacteria."
		RT Arch. Microbiol. 152:335-341(1989);
		RL -!- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon dioxide.
		CC -!- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced acceptor.
		CC -!- COFACTOR: Molybdenum (molybdopterin).
		CC -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND SMALL.
		DR PIR: PL0140; PL0140.
		KW Oxidoreductase; Molybdenum.
		FT NON-TER 4 4 MW;
		SQ SEQUENCE 4 AA; 441 MW; 7761E876F0000000 CRC64;
		Query Match 0.0%; Score 0; DB 1; Length 4;
		Best Local Similarity 0.0%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;
		Matches 0; Conservative 0; Mismatches 1; Indels 1; Gaps 0;
	Qy 1 X 1	QY 1 X 1
	Db 1 M 1	Db 3 S 3
	RESULT 6 DCMS_PSECH STANDARD; PRT; 4 AA.	DE /FTID=VAR_005201. 6B05B862A0000000 CRC64;
	ID DCMS_PSECH STANDARD; PRT; 4 AA.	RN
	AC P19918;	RT
		DT 01-FEB-1991 (Rel. 17, Created)
		DT 01-FEB-1991 (Rel. 17, Last sequence update)
		DT 15-JUN-2002 (Rel. 41, Last annotation update)
		DE Carbon monoxide dehydrogenase small chain (EC 1.2.99.2) (CO



Best Local Similarity 0.0%; Pred. No. 0; Matches 0; Conservative 0; Mismatches 0; Indels 1; Gaps 0; O:

Qy 1 X 1  
Db 1 F 1

**RESULT 12**

ID	FLRN_ANTEL	STANDARD;	PRT;	4 AA.
AC	P58707;			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Antho RNamide.			
VS	Anthopleura elegantissima (Sea anemone).			
OC	Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;			
OC	Nyantheae; Actiniidae; Anthopleura.			
OX	NCBI_TAXID=6110;			
RP	SEQUENCE, AND MASS-SPECTROMETRY.			
RX	GRIMMELIKHUIZEN C.J.P., RINEHART K.L. JR., GRAFF D., REINSCHEID R.K., NOTHAAKER H.-P., STALEY A.L.; "Isolation of L-3-phenylalanyl-leu-Arg-Asn-NH2 (Antho-Rnamide), a sea anemone neuropeptide containing an unusual amino-terminal blocking group.", Proc. Natl. Acad. Sci. U.S.A. 87:5410-5414 (1990).			
CC	-!- SUBCELLULAR LOCATION: Secreted.			
CC	-!- TISSUE SPECIFICITY: Neuron specific.			
KW	Neuropeptide; Amidation.			
FT	MOD_RES 1 1 L-3-PHENYLALACTYL.			
FT	MOD_RES 4 4 AMIDATION.			
SQ	SEQUENCE 4 AA; 549 MW; 64540729A0000000 CRC64;			

Query Match 0.0%; Score 0; DB 1; Length 4;  
Best Local Similarity 0.0%; Pred. No. 0; Matches 0; Indels 0; Gaps 0; O:

Qy 1 X 1  
Db 1 F 1

**RESULT 13**

ID	FMRF_MACN1	STANDARD;	PRT;	4 AA.
AC	P01162;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUN-1998 (Rel. 36, Last annotation update)			
DE	FMRamide (Peak C) (Cardioexcitatory neuropeptide).			
OS	Macrocallista nimbosa (Sun-ray clam), Nereis virens (Sandworm), Hirudo medicinalis (Medicinal leech), and Helisoma trivolvis (snail).			
OS	Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroida; Veneroidea; Veneridae; Macrocallista.			
OX	NCBI_TAXID=6594, 6353, 6421, 27815;			
RP	SEQUENCE, AND SYNTHESIS.			
RC	SPECIES=M. nimbosa; TISSUE=Cerebral pedal, and Visceral ganglion;			
RA	Price D.A., Greenberg M.J.;			
RA	"Structure of a molluscan cardioexcitatory neuropeptide.", Science 197:670-671 (1977).			
RA	SEQUENCE, AND CHARACTERIZATION.			
RC	SPECIES=M. nimbosa; TISSUE=Ganglion;			
RX	MEDLINE=78012038; PubMed=909875;			
RA	Price D.A., Greenberg M.J.;			

Best Local Similarity 0.0%; Pred. No. 0; Matches 0; Conservative 0; Mismatches 0; Indels 1; Gaps 0; O:

Qy 1 X 1  
Db 1 F 1

**RESULT 14**

ID	FYRL_ANTEL	STANDARD;	PRT;	4 AA.
AC	P38706;			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Antho-Riamide I (Contains Antho-Riamide II).			
OS	Anthopleura elegantissima (Sea anemone).			
OC	Eukaryota; Metazoa; Chidaria; Anthozoa; Zoantharia; Actiniaria;			
OX	NCBI_TAXID=6110;			
RN	[1]			
RP	SEQUENCE.			
RX	MEDLINE=92270459; PubMed=18210967;			
RA	Nothacker H.-P., Rinehart K.L. Jr., McFarlane I.D., Grimmelikhuizen C.J.P.;			
RA	"Isolation of two novel neuropeptides from sea anemones: the unusual, biologically active L-3-phenylalactyl-Tyr-Arg-Ile-NH2 and its des-phenylalactyl fragment Tyr-Arg-Ile-NH2.", Peptides 12:1165-1173 (1991).			
RA	Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188 (1993).			
RP	FUNCTION.			
RX	MEDLINE=93391436; PubMed=8397415;			
RA	McFarlane I.D., Hudman D., Notthacker H.-P., Grimmelikhuizen C.J.P.;			
RT	"The expansion behaviour of sea anemones may be coordinated by two inhibitory neuropeptides, Antho-Riamide and Antho-Riamide.", Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188 (1993).			
RL				

CC    -!- FUNCTION: Inhibits spontaneous contractions in several muscle groups. May be involved in the expansion phase of feeding behaviour in sea anemones.

CC    -!- SUBCELLULAR LOCATION: Secreted.

CC    -!- TISSUE SPECIFICITY: Neuron-specific.

DR    InterPro: IPR001023; Hsp70.

KW    Neuropeptide; Amidation.

FT    CHAIN    1    ANTHO-RIAMIDE I.

FT    CHAIN    2    ANTHO-RIAMIDE II.

FT    MOD\_RES    1    L-3-PHENYLALACTYL.

FT    MOD\_RES    4    AMIDATION.

SO    SEQUENCE    4 AA;    598 MW;    60441B59A000000 CRC64;

Query Match    0.0%;    Score 0;    DB 1;    Length 4;

Best Local Similarity    0.0%;    Pred. No. 0;

Matches    0;    Conservative 0;    Mismatches 1;    Indels 0;    Gaps 0;

Qy    1 x 1

Db    1 F 1

---

RESULT 15

SPLOCTMI	OCP1_OCTMI	STANDARD	PRT	4 AA.
2	P58648;			
DT	15-JUN-2002	(Rel. 41, Created)		
DT	15-JUN-2002	(Rel. 41, Last sequence update)		
DT	15-JUN-2002	(Rel. 41, Last annotation update)		
DE	C cardioactive peptides Ocp-1/Ocp-2.			
OS	Octopus minor (Octopus)			
OC	Eukaryota; Mollusca; Coleoidea; Octopoda; Incirrata; Octopodidae; Octopus; NCBITaxID=89766;			
RN	[1]			
RP	SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.			
RC	TISSUE=Brain;			
RX	MEDLINE=20336815; PubMed=10876044;			
RA	"Cardioactive peptides isolated from the brain of a Japanese octopus, Iwakoshi E., Hisada M., Minakata H.;"			
RT	Octopus minor."			
RL	Peptides 21:623-630(2000).			
CC	-!- FUNCTION: Cardioactive; has both positive chronotropic and inotropic effects on the heart. Ocp-2 is a 1000 time less active than Ocp-1.			
CC	-!- SUBCELLULAR LOCATION: Secreted.			
CC	-!- PTM: Ocp-2 has L-Phe instead of D-Phe.			
CC	-!- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.			
KW	Hormone; D-amino acid.			
FT	MOD_RES    2    D-PHENYLALANINE.			
SO	SEQUENCE    4 AA;    394 MW;    6AA879C810000000 CRC64;			
Query Match    0.0%;    Score 0;    DB 1;    Length 4;				
Best Local Similarity    0.0%;    Pred. No. 0;				
Matches    0;    Conservative 0;    Mismatches 1;    Indels 0;    Gaps 0;				
Qy    1 x 1				
Db    1 G 1				

GenCore version 5.1.3	Q15897	homo sapien	7	4
Copyright (c) 1993 - 2003 Compugen Ltd.	Q15903	homo sapien	7	4
OM protein - protein search, using sw model	P83274	macrobrachii	0	0
Run on : February 12, 2003, 11:42:45 ; Search time 60.2667 Seconds (without alignments) 27.351 Million cell updates/sec	P28742	oryctolegus	7	5
Title: US-09-660-302C-1	P92372	haynalia v	0	0
Perfect score: 8	P92403	lophopyrum	0	0
Sequence: 1 XXXXXXXX 8	P92425	pseudoergen	0	0
Scoring table: BLOSUM62	P92427	herdmania p	0	0
Gapop 10.0 , Gapext 0.5	P92390	peridictyon	0	0
Searched: 671580 seqs, 20604715 residues	P92214	heteranthel	0	0
Total number of hits satisfying chosen parameters: 671580	P92226	critopopsis	0	0
Minimum DB seq length: 0	P92214	amblyopyrum	0	0
Maximum DB seq length: 2000000000	P92430	aegilops ta	0	0
Post-processing: Minimum Match 0%	P92221	brunus linear	0	0
Maximum Match 100%	P92442	taeniatheru	0	0
Listing first 45 summaries	P92381	hordeum bra	0	0
Database : SPTREMBL_21:	P92214	hordeum vul	0	0
1: sp_archaea:*	P92218	australyorum	0	0
2: sp_bacteria:*	P92440	thinopyrum	0	0
3: sp_fungi:*	P92210	agropyron c	0	0
4: sp_human:*	P99182	gnatholebia	0	0
5: sp_invertebrate:*	Q95945	saccharonyc	0	0
6: sp_mammal:*	Q98866	spinacia ol	0	0
7: sp_mhc:*	P92393	psathyroa	0	0
8: sp_organelle:*	P92218	hordeum mar	0	0
9: sp_phage:*	P92421	mar	0	0
10: sp_plant:*	P92385	hordeum	0	0
11: sp_rabbit:*	P92210	agropyron max	0	0
12: sp_virus:*	Q92223	Q9c5b5; arabidopsis	0	0
13: sp_vertebrate:*	P93233	lycopersico	0	0
14: sp_unclassified:*	P82445	nicotiana t	0	0
15: sp_rfivirus:*	0	0	0	0
16: sp_bacterioplasm:*	0	0	0	0
17: sp_archaea:*	0	0	0	0
		ALIGMENTS		
		RESULT 1		
	Q08433			
	ID Q08433	PRELIMINARY;	PRT;	4 AA.
	AC 008433;			
	DT 01-NOV-1996 (TREMBLrel. 01, Created)			
	DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
	DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)			
	DE UDP-glucuronosyltransferase, microsomal (EC 2.4.1.17) (UDPGT)			
	DE (Fragment)			
	DE Rattus norvegicus (Rat).			
	OS Bukarvota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
	OC NCBI_TaxID=10116;			
	RN [1]	SEQUENCE FROM N.A.		
	RP STRAIN=GUNN;			
	RC MEDLINE#91382758; PubMed=18404866;			
	RA Sato H, Aono S, Kashiiwana S, Koiwai O;			
	RT Genetic defect of bilirubin UDP-glucuronosyltransferase in the			
	RT hyperbilirubinemic Gunn rat.;"			
	RL Biochem Biophys Res Commun. 177:1161-1164(1991)			
	CC -1 FUNCTION: UDPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND			
	CC SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND			
	CC ENDOGENOUS COMPOUNDS.			
	CC -1 CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR			
	CC BETA-D-GLUCORONOSIDE.			
	CC -1 SUBCELLULAR LOCATION: MICROSOME.			
	DR EMBL; S8636; ARB1929.1;			
	KW Transf erase; Glycosyltransferase; Microsome; Multigene family.			
	CC ENDONUCLEOTIDE GLYCOSYLTRANSFERASES.			
	FT NON-TER FT NON-TER	Score 0; DB 11; Length 4;		
	FT NON-TER	Pred. No. 0;		
	SQ SEQUENCE 4 AA: 473 MW: 633732C420000000 CRC64;	Mismatches 0;保守型 1; Indels 0; Gaps 0;		
	Query Match Best Local Similarity 0.0%; Matches 0; Conservative 0;			
	QY 1 x 1			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score distribution, and is derived by analysis of the result being printed.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	0	0.0	4	11	Q08433	rattus norvegicus
2	0	0.0	5	2	P83073	bacillus ce
3	0	0.0	5	10	Q99007	hordeum vul
4	0	0.0	5	13	P83308	gallus gallus
5	0	0.0	6	10	P82181	spinacia ol
6	0	0.0	6	10	P82541	spinacia ol
7	0	0.0	6	10	P82182	spinacia ol
8	0	0.0	7	2	007354	synchococcus
9	0	0.0	7	2	047029	enterobacte
10	0	0.0	7	2	050556	actinobacil
11	0	0.0	7	2	034028	sphingomonas
12	0	0.0	7	2	047477	escherichia
13	0	0.0	7	2	047505	047505
14	0	0.0	7	2	P70804	azotobacter
15	0	0.0	7	2	054248	streptomyce
16	0	0.0	7	2	P72081	nocardia la

Db	1 N 1	RESULT 2 PB3073 PRELIMINARY; PRT; 5 AA. ID PB3073; AC P83073; DT 01-OCT-2001 (TREMBLrel. 18, Created) DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update) DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update) DE 88 kDa protein (fragment). OS Bacillus cereus. OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales; OC Bacillaceae; Bacillus. NCBI_TaxID=1396; RN [1]. RP SEQUENCE. RC STRAIN=NCIMB 11796; RA Browne N.; Dowds B.C.A.; FT Submitted (JUL-2001) to the SWISS-PROT data bank. SQ SEQUENCE 5 AA; 623 MW; 6B01AAA336F00000 CRR64; Query Match Score: 0; DB 2; Length: 5; Best Local Similarity 0.0%; Pred. No. 0; Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0; QY 1 X 1 Db 4 T 4	RESULT 3 Q99007 PRELIMINARY; PRT; 5 AA. ID Q99007; AC Q99007; DT 01-NOV-1996 (TREMBLrel. 01, Created) DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update) DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update) DE Alpha-amylase (EC 3.2.1.1) (Fragment). GN AMY1. OS Hordeum vulgare (Barley). OC Viridiplanteae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Pooidae; Triticeae; Hordeum. NCBI_TaxID=4513; RN RP SEQUENCE FROM N.A. TISSUE=ALEURONE LAYER; RC STRAIN=HIMALAYA; PUBMED=91329704; RX PUBLISHER=Pubmed; PUBLISHED=1831055; [1] Jacobson J.V., Close T.J.; "Control of transient expression of chimaeric genes by gibberellic acid and abscisic acid in protoplasts prepared from mature barley aleurone layers.", Plant Mol. Biol. 16: 713-721(1991). CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES. CC -!- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY. CC -!- MISCELLANEOUS: THERE ARE AT LEAST 4 TYPES OF ALPHA-AMYLASE IN BARLEY. EMBL: X54643; CAA38455; 1; DR Hydrolase; Glycosidase; Carbohydrate metabolism; Seed; Germination; Calcium; Multigene family. KW Non-TER 5 AA; 600 MW; 61E3344DD6F00000 CRR64; SQ SEQUENCE 5 AA; 600 MW; 61E3344DD6F00000 CRR64;	RESULT 4 PB3308 PRELIMINARY; PRT; 5 AA. ID PB3308; AC P83308; DT 01-JUN-2002 (TREMBLrel. 21, Created) DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update) DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update) DE FMRFamide-like neuropeptide (LPRLF-anide). OS Gallus gallus (Chicken). OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; OC Gallo. NCBITaxonID=9031; RN [1]. RP SEQUENCE, AND SYNTHESIS. RC TISSUE-BRAIN; RT PubMed=6137771; RA Dockray G.J.; Reeve J.R. Jr.; Shively J.; Gayton R.J.; Barnard C.S.; "A novel active pentapeptide from chicken brain identified by antibodies to FMRFamide."; RT Nature 305:328-330(1983). RL CC -!- FUNCTION: MAY FUNCTION AS A NEUROTRANSMITTER OR MODULATOR. CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE) FAMILY. KW Neuropeptide. SQ SEQUENCE 5 AA; 645 MW; 69D4073767400000 CRC64; Query Match Score: 0.0%; Best Local Similarity 0.0%; Pred. No. 0; Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db	1 X 1	RESULT 5 P82181 PRELIMINARY; PRT; 6 AA. ID P82181; AC P82181; DT 01-JUN-2000 (TREMBLrel. 14, Created) DT 01-JUN-2000 (TREMBLrel. 14, Last sequence update) DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update) DE Chloroplast 50S ribosomal protein L10 beta (Fragment). OS Spinacia oleracea (Spinach). OC Eukaryota; Viridiplanteae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eu dicots; core eudicots; CC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia. NCBITaxonID=5562; OX [1]. RN RP SEQUENCE TO THE 2D-GEL IT IS 16.5 KDA. RC STRAIN=CV. ALWARO; TISSUE=LEAF; RX MEDLINE=20435798; PubMed=10874046; RA Yamaguchi K.; Subramanian A.R.; RT "The plastid ribosomal proteins. Identification of all the proteins in the 50 S subunit of an organelle ribosome (chloroplast)." ; RL J. Biol. Chem. 275: 28466-28482(2000). CC -!- SUBCELLULAR LOCATION: CHLOROPLAST. CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES. CC -!- MISCELLANEOUS: ON THE 2D-GEL IT IS 16.5 KDA. CC -!- SIMILARITY: BELONGS TO THE 2D-GEL IT IS 16.5 KDA. DR InterPro; IPR012190; Ribosomal_L10_110. DR Pfam; PF00466; Ribosomal_L10; PARTIAL. DR PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL. KW Ribosomal protein; Chloroplast; rRNA-binding. FT NON-TER 6 AA; 667 MW; 6321BA15B05DB000 CRC64; SQ Sequence 6 AA; 667 MW; 6321BA15B05DB000 CRC64; Query Match Score: 0.0%; Best Local Similarity 0.0%; Pred. No. 0; Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	RESULT 5 P82181 PRELIMINARY; PRT; 6 AA. ID P82181; AC P82181; DT 01-JUN-2000 (TREMBLrel. 14, Created) DT 01-JUN-2000 (TREMBLrel. 14, Last sequence update) DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update) DE Chloroplast 50S ribosomal protein L10 beta (Fragment). OS Spinacia oleracea (Spinach). OC Eukaryota; Viridiplanteae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eu dicots; core eudicots; CC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia. NCBITaxonID=5562; OX [1]. RN RP SEQUENCE TO THE 2D-GEL IT IS 16.5 KDA. RC STRAIN=CV. ALWARO; TISSUE=LEAF; RX MEDLINE=20435798; PubMed=10874046; RA Yamaguchi K.; Subramanian A.R.; RT "The plastid ribosomal proteins. Identification of all the proteins in the 50 S subunit of an organelle ribosome (chloroplast)." ; RL J. Biol. Chem. 275: 28466-28482(2000). CC -!- SUBCELLULAR LOCATION: CHLOROPLAST. CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES. CC -!- MISCELLANEOUS: ON THE 2D-GEL IT IS 16.5 KDA. CC -!- SIMILARITY: BELONGS TO THE 2D-GEL IT IS 16.5 KDA. DR InterPro; IPR012190; Ribosomal_L10_110. DR Pfam; PF00466; Ribosomal_L10; PARTIAL. DR PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL. KW Ribosomal protein; Chloroplast; rRNA-binding. FT NON-TER 6 AA; 667 MW; 6321BA15B05DB000 CRC64; SQ Sequence 6 AA; 667 MW; 6321BA15B05DB000 CRC64; Query Match Score: 0.0%; Best Local Similarity 0.0%; Pred. No. 0; Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
Db	2 A 2	RESULT 5 P82181 PRELIMINARY; PRT; 6 AA. ID P82181; AC P82181; DT 01-JUN-2000 (TREMBLrel. 14, Created) DT 01-JUN-2000 (TREMBLrel. 14, Last sequence update) DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update) DE Chloroplast 50S ribosomal protein L10 beta (Fragment). OS Spinacia oleracea (Spinach). OC Eukaryota; Viridiplanteae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eu dicots; core eudicots; CC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia. NCBITaxonID=5562; OX [1]. RN RP SEQUENCE TO THE 2D-GEL IT IS 16.5 KDA. RC STRAIN=CV. ALWARO; TISSUE=LEAF; RX MEDLINE=20435798; PubMed=10874046; RA Yamaguchi K.; Subramanian A.R.; RT "The plastid ribosomal proteins. Identification of all the proteins in the 50 S subunit of an organelle ribosome (chloroplast)." ; RL J. Biol. Chem. 275: 28466-28482(2000). CC -!- SUBCELLULAR LOCATION: CHLOROPLAST. CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES. CC -!- MISCELLANEOUS: ON THE 2D-GEL IT IS 16.5 KDA. CC -!- SIMILARITY: BELONGS TO THE 2D-GEL IT IS 16.5 KDA. DR InterPro; IPR012190; Ribosomal_L10_110. DR Pfam; PF00466; Ribosomal_L10; PARTIAL. DR PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL. KW Ribosomal protein; Chloroplast; rRNA-binding. FT NON-TER 6 AA; 667 MW; 6321BA15B05DB000 CRC64; SQ Sequence 6 AA; 667 MW; 6321BA15B05DB000 CRC64; Query Match Score: 0.0%; Best Local Similarity 0.0%; Pred. No. 0; Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	RESULT 5 P82181 PRELIMINARY; PRT; 6 AA. ID P82181; AC P82181; DT 01-JUN-2000 (TREMBLrel. 14, Created) DT 01-JUN-2000 (TREMBLrel. 14, Last sequence update) DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update) DE Chloroplast 50S ribosomal protein L10 beta (Fragment). OS Spinacia oleracea (Spinach). OC Eukaryota; Viridiplanteae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eu dicots; core eudicots; CC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia. NCBITaxonID=5562; OX [1]. RN RP SEQUENCE TO THE 2D-GEL IT IS 16.5 KDA. RC STRAIN=CV. ALWARO; TISSUE=LEAF; RX MEDLINE=20435798; PubMed=10874046; RA Yamaguchi K.; Subramanian A.R.; RT "The plastid ribosomal proteins. Identification of all the proteins in the 50 S subunit of an organelle ribosome (chloroplast)." ; RL J. Biol. Chem. 275: 28466-28482(2000). CC -!- SUBCELLULAR LOCATION: CHLOROPLAST. CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES. CC -!- MISCELLANEOUS: ON THE 2D-GEL IT IS 16.5 KDA. CC -!- SIMILARITY: BELONGS TO THE 2D-GEL IT IS 16.5 KDA. DR InterPro; IPR012190; Ribosomal_L10_110. DR Pfam; PF00466; Ribosomal_L10; PARTIAL. DR PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL. KW Ribosomal protein; Chloroplast; rRNA-binding. FT NON-TER 6 AA; 667 MW; 6321BA15B05DB000 CRC64; SQ Sequence 6 AA; 667 MW; 6321BA15B05DB000 CRC64; Query Match Score: 0.0%; Best Local Similarity 0.0%; Pred. No. 0; Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	

Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 A 1

**RESULT 6**  
P82541 PRELIMINARY; PRT; 6 AA.  
ID P82541;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-JUN-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE Chloroplast 30S ribosomal protein S19 beta (Fragment).  
KW Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; euodiatelyoids; core eudicots;  
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.  
OX NCBI\_TAXID=3562;  
RN [1]  
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.  
RC STRAIN=CV\_ALNPRO; TISSUE=LEAF;  
RX MEDLINE=20435797; PubMed=1084039;  
RA Yamauchi K., von Knoblauch K., Subramanian A.R.;  
RT "The plastid ribosomal proteins. Identification of all the proteins in  
the small subunit of an organelle ribosome (chloroplast).";  
RL J. Biol. Chem. 37:28455-28465(2000).  
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA.  
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
CC -!- MASS SPECTROMETRY: MW=10477.0; METHOD=ELECTROSPRAY.  
CC -!- MASS SPECTROMETRY: MW=10495.0; METHOD=MALDI.  
CC -!- MISCELLANEOUS: S19 ALPHA AND BETA FORMS DIFFER IN PI. S19 BETA  
FORM IS THE MINOR BASIC FORM.  
CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 1.2 kDa.  
CC -!- MISCELLANEOUS: BELONGS TO THE S19 FAMILY OF RIBOSOMAL PROTEINS.  
DR InterPro:IPR002222; Ribosomal\_S19.  
DR Pfam:PF00203; Ribosomal\_S19; PARTIAL.  
DR PRINTS: PR00975; RIBOSOMALS19; PARTIAL.  
DR PROSITE: PS00333; RIBOSOMAL\_S19; PARTIAL.  
KW Ribosomal protein; Chloroplast; rRNA-binding.

NON\_TER 6  
SEQUENCE 6 AA; 732 MW; 63333375A411C000 CRC64;  
Query Match 0.0%; Score 0; DB 10; Length 6;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 T 1

**RESULT 7**  
P82182 PRELIMINARY; PRT; 6 AA.  
ID P82182;  
AC P82182;  
DT 01-JUN-2000 (TREMBLrel. 14, Created)  
DT 01-JUN-2000 (TREMBLrel. 14, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Chloroplast 50S ribosomal protein L10 gamma (Fragment).  
OS Spinacia oleracea (Spinach).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; euodiatelyoids; core eudicots;  
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.  
OX NCBI\_TAXID=3562;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=CV\_ALVARO; TISSUE=LEAF;  
RX MEDLINE=20435798; PubMed=10874046;

RA Yamaguchi K., Subramanian A.R.;  
RT "The plastid ribosomal proteins. Identification of all the proteins in  
the 50 S subunit of an organelle ribosome (chloroplast).";  
RL J. Biol. Chem. 275:28466-28484(2000).  
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.  
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 kDa.  
CC -!- MISCELLANEOUS: BELONGS TO THE L10 FAMILY OF RIBOSOMAL PROTEINS.  
DR InterPro:IPR011790; Ribosomal\_L10.  
DR InterPro:IPR023633; Ribosomal\_L10sub.  
DR Pfam:PF00466; Ribosomal\_L10; PARTIAL.  
DR PROSITE: PS01109; RIBOSOMAL\_L10; PARTIAL.  
KW Ribosomal protein; Chloroplast; rRNA-binding.  
NON\_TER 6  
SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;  
Query Match 0.0%; Score 0; DB 10; Length 6;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 X 1  
Db 1 A 1

**RESULT 8**  
P07354 PRELIMINARY; PRT;  
ID P07354;  
AC 007354;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE NIFK (Fragment).  
GN OS  
RT Synechococcus sp. (strain PCC 8801 / RF-1) (Cyanothecae PCC 8801).  
OC Bacteria; Cyanobacteria; Chroococcales; Cyanothecae.  
OX NCBI\_TAXID=41431;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=90231861; PubMed=10217509;  
RA Huang T.C.; Lin R.F.; Chu M.K.; Chen H.M.;  
RT "Organization and expression of nitrogen-fixation genes in the aerobic  
nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain  
RT RF-1";  
RL Microbiology 145:743-753(1999).  
DR EMBL; AF003700; AAC35193.1; -.  
NON\_TER 1  
SEQUENCE 7 AA; 849 MW; 7412C7AA9DSB030 CRC64;  
Query Match 0.0%; Score 0; DB 2; Length 7;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 X 1  
Db 2 S 2

**RESULT 9**  
Q47029 PRELIMINARY; PRT;  
ID Q47029;  
AC 047029;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
DE Ad A1 protein (Fragment).  
GN AD A1.  
OS Enterobacter cloacae.  
OC Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Enterobacter.  
OX NCBI\_TAXID=550;

[1]	RN SEQUENCE FROM N.A. RX MEDLINE=94079349; PubMed=8257126; RA Rather P.N.; Mann P.A.; Mierwa R.; Hare R.S.; Miller G.H.; Shaw K.J.; RT "Analysis of the aac(3)-VIA gene encoding a novel 3-N-acetyltransferase." RL Antimicrob Agents Chemother. 37:2074-2079(1993). DR M88012; AAA16193.1; -. FT NON_TER 1 1 SQ SEQUENCE 7 AA; 744 MW; 63386226321A030 CRC64; Query Match 0.0%; Score 0; DB 2; Length 7; Best Local Similarity 0.0%; Pred. No. 0; Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0; Qy 1 X 1 Db 2 T 2	RA Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases. RL EMBL; U88298; AAB66311.1; KW Dioxygenase. RT NON_TER 1 SQ SEQUENCE 7 AA; 868 MW; 71A452D1A699D460 CRC64;
	RESULT 12 Q4477 ID Q4477 PRELIMINARY; PRT; 7 AA. AC 047477; PRT; DT 01-NOV-1996 (TREMBLrel. 01, Created) DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update) DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update) DE Tpi protein (Fragment). GN TPI. OS Escherichia coli. OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; OC Escherichia. NCBI_TaxID=562; OX RN SEQUENCE OF 7-7 FROM N.A. RP MEDLINE=85203917; PubMed=3158524; RX Hellinga H.W., Evans P.R.; RA "Nucleotide sequence and high-level expression of the major RT Escherichia coli phosphofructokinase."; RT Eur. J. Biochem. 149:365-373(1985). RN [2] RP SEQUENCE FROM N.A. RN SEQUENCE FROM N.A. RA Evans P.; RA Hellinga H.W., Evans P.R.; RA Submitted (OCT-1986) to the EMBL/GenBank/DDBJ databases. DR X02519; CAA26359.1; -. FT NON_TER 1 1 SQ SEQUENCE 7 AA; 773 MW; 7416D3DDDB1DB0 CRC64;	RA Kraig E.; RT "cis Elements and trans factors are both important in strain-specific RT regulation of the leukotoxin gene in Actinobacillus RT actinomycetemcomitans"; RT Infect. Immun. 64:3451-3460(1996). RL DR EMBL; U51862; AAB88721.1; -. FT NON_TER 1 1 SQ SEQUENCE 7 AA; 832 MW; 6DCB42D767340420 CRC64;
	Query Match 0.0%; Score 0; DB 2; Length 7; Best Local Similarity 0.0%; Pred. No. 0; Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0; Qy 1 X 1 Db 1 A 1	Query Match 0.0%; Score 0; DB 2; Length 7; Best Local Similarity 0.0%; Pred. No. 0; Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0; Qy 1 X 1 Db 1 A 1
	RESULT 13 Q47505 ID Q47505 PRELIMINARY; PRT; 7 AA. AC 047505; PRT; DT 01-NOV-1996 (TREMBLrel. 01, Created) DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update) DE McCa protein. GN Escherichia coli. OS Plasmid pMCCC7. OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; OC Escherichia. NCBI_TaxID=562; OX RN SEQUENCE FROM N.A. RP MEDLINE=96093297; PubMed=8522520; RX Gonzalez-Pastor J.E., San Millan J.L., Castilla M.A., Moreno F.; RT "Structure and organization of plasmid genes required to produce the RT translation inhibitor microcin C7;"	RA Sphingomonas chungbukensis. OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae; OC Sphingomonas. OX NCBI_TaxID=56193; RN [1] RP SEQUENCE FROM N.A. RC STRAIN=DJ77;

RL J. Bacteriol. 177:7131-7140(1995).  
 KW EMBL; X57583; CAA40808.1; -.  
 SQ SEQUENCE 7 AA; 763 MW; 644DD44861B406F0 CRC64;  
 Query Match 0.0%; Score 0; DB 2; Length 7;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0;  
 Gaps 0;  
 QY 1 X 1  
 Db 3 T 3

---

RESULT 14  
 P70804 ID P70804 PRELIMINARY; PRT; 7 AA.  
 IC P70804; 01-FEB-1997 (TREMBLrel. 02, Created)  
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE AlgT protein (Fragment).  
 GN ALGT.  
 OS Azotobacter vinelandii.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 Azotobacter.  
 OX NCBI\_TaxID=54;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=N;  
 RX MEDLINE=96427318; PubMed=8830682;  
 RA Rehman, B.H.A., Ertesvag H., Valla S.;  
 RT "A new Azotobacter vinelandii mannuronan C-5-epimerase gene (algG) is  
 part of an alg gene cluster physically organized in a manner similar  
 to that in *Pseudomonas aeruginosa*."  
 RT J. Bacteriol. 178:5884-5889(1996).  
 DR EMBL; X87973; CAA61230.1;  
 FT NON-TER 1 1  
 SQ SEQUENCE 7 AA; 684 MW; 71B5A5A2D1AED0 CRC64;  
 Query Match 0.0%; Score 0; DB 2; Length 7;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0;  
 Gaps 0;  
 QY 1 X 1  
 Db 2 T 2

---

RESULT 15  
 Q54248 ID Q54248 PRELIMINARY; PRT; 7 AA.  
 AC Q54248;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE RplO protein (Fragment).  
 GN RPLO.  
 OS Streptomyces griseus.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinomycetidae; Streptomycetaceae; Streptomyces.  
 NCBI\_TaxID=1911;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=N2-3-11;  
 RX MEDLINE=200111291; PubMed=10542330;  
 RA Poehling S., Piepersberg W., Weimer U.F.;  
 RT "Analysis and regulation of the sec Y gene from *Streptomyces griseus*  
 N2-3-11 and interaction of the SEC Y protein with the SecA protein.";  
 RL Blochim. Biophys. Acta 1447:298-302(1999).  
 DR EMBL; X95915; CAA65160.1; -.  
 FT NON-TER 1 1  
 SQ SEQUENCE 7 AA; 760 MW; 72C72B01B2D1B2A0 CRC64;

**THIS PAGE BLANK (USPTO).**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compigen Ltd.

OM protein - protein search, using sw model  
Run on: February 12, 2003, 11:39:30 ; Search time 28 Seconds  
(without alignments)  
33.313 Million cell updates/sec

Title: US-09-660-302C-7  
Perfect score: 43  
Sequence: 1 CEEDEFYR 7

Scoring table: BLOSUM62  
Gapox 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470 ✓

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002:\*

1: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1980.DAT:\*

2: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1981.DAT:\*

3: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1982.DAT:\*

4: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1983.DAT:\*

5: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1984.DAT:\*

6: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1985.DAT:\*

7: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1986.DAT:\*

8: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1987.DAT:\*

9: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1988.DAT:\*

10: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1989.DAT:\*

11: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1990.DAT:\*

12: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1991.DAT:\*

13: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1992.DAT:\*

14: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1993.DAT:\*

15: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1994.DAT:\*

16: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1995.DAT:\*

17: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1996.DAT:\*

18: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1997.DAT:\*

19: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1998.DAT:\*

20: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA1999.DAT:\*

21: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA2000.DAT:\*

22: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA2001.DAT:\*

23: /SIDS2/gcdata/geneseq/geneseqp-emb1/AA2002.DAT:\*

11 38 88.4 340 23 AU75496  
12 38 88.4 637 10 AP92108  
13 38 88.4 638 9 AA81326  
14 38 88.4 638 19 AW33394  
15 38 88.4 648 22 ABB11437  
16 36 83.7 84 20 AY30432  
17 36 83.7 94 21 ABB15317  
18 36 83.7 91 17 AR91701  
19 36 83.7 91 20 AYX30393  
20 36 83.7 91 20 AYX30454  
21 36 83.7 91 21 AB15346  
22 36 83.7 404 19 AW54077  
23 36 83.7 406 22 AB95251  
24 36 83.7 423 19 AW54078  
25 34 79.1 49 22 AO11680  
26 34 79.1 102 23 ABP30709  
27 34 79.1 615 22 ABG07365  
28 34 79.1 1589 22 AU00294  
29 34 79.1 1591 22 AB32113  
30 34 79.1 1591 22 ABM30819  
31 34 79.1 1591 22 ABM5938  
32 34 79.1 1591 22 ABP22654  
33 34 79.1 1591 22 AM58054  
34 34 79.1 1591 22 AM18331  
35 34 79.1 1591 22 AM30819  
36 34 79.1 1591 22 ABG40136  
37 34 79.1 1591 23 ABG40136  
38 33 76.7 75 20 AYX30414  
39 33 76.7 1591 22 AM70495  
40 33 76.7 1591 22 AM18331  
41 33 76.7 78 20 AYX30420  
42 33 76.7 78 21 AB15305  
43 33 76.7 82 20 AYX30399  
44 33 76.7 82 21 AB15293  
45 33 76.7 82 21 AB15307

## ALIGNMENTS

RESULT 1  
AY32791  
ID AY32791 standard; peptide: 7 AA.  
XX  
AC AY32791;  
XX  
DT 09-NOV-1999 (first entry)  
XX  
DE Proteolytic cleavage signal site used in inhibiting receptor proteolysis.  
XX  
KW Signal transduction; Proteolytic cleavage; cleavage signal site;  
KW Ubiquitin; proteosome binding site; muscle wasting; renal tubular defect;  
KW uraemia; diabetes; Cushing's disease; eating disorder; AIDS;  
KW growth hormone deficiency.  
XX  
OS Mammalia.  
XX  
PN EP943624-A1.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	100.0	7	AAY32791	Proteolytic cleavage
2	38	88.4	33	AAR2156	Plasmid pRSR8-3 en
3	38	88.4	246	AAR56389	Human growth hormo
4	38	88.4	249	AAR06367	Hormone binding re
5	38	88.4	269	AAR05045	Soluble human grow
6	38	88.4	269	AAR10426	Human somatogenic
7	38	88.4	269	AAY31767	Human soluble grow
8	38	88.4	269	AAYW82802	Human soluble soma
9	38	88.4	269	AAY78429	Soluble part of th
10	38	88.4	315	AAU75499	Human fusion prote

Pred. No. is the number of results predicted by chance to have a a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

1	43	100.0	7	AAY32791	Proteolytic cleavage
2	38	88.4	33	AAR2156	Plasmid pRSR8-3 en
3	38	88.4	246	AAR56389	Human growth hormo
4	38	88.4	249	AAR06367	Hormone binding re
5	38	88.4	269	AAR05045	Soluble human grow
6	38	88.4	269	AAR10426	Human somatogenic
7	38	88.4	269	AAY31767	Human soluble grow
8	38	88.4	269	AAYW82802	Human soluble soma
9	38	88.4	269	AAY78429	Soluble part of th
10	38	88.4	315	AAU75499	Human fusion prote

1	43	100.0	7	AAY32791	Proteolytic cleavage
2	38	88.4	33	AAR2156	Plasmid pRSR8-3 en
3	38	88.4	246	AAR56389	Human growth hormo
4	38	88.4	249	AAR06367	Hormone binding re
5	38	88.4	269	AAR05045	Soluble human grow
6	38	88.4	269	AAR10426	Human somatogenic
7	38	88.4	269	AAY31767	Human soluble grow
8	38	88.4	269	AAYW82802	Human soluble soma
9	38	88.4	269	AAY78429	Soluble part of th
10	38	88.4	315	AAU75499	Human fusion prote

1	43	100.0	7	AAY32791	Proteolytic cleavage
2	38	88.4	33	AAR2156	Plasmid pRSR8-3 en
3	38	88.4	246	AAR56389	Human growth hormo
4	38	88.4	249	AAR06367	Hormone binding re
5	38	88.4	269	AAR05045	Soluble human grow
6	38	88.4	269	AAR10426	Human somatogenic
7	38	88.4	269	AAY31767	Human soluble grow
8	38	88.4	269	AAYW82802	Human soluble soma
9	38	88.4	269	AAY78429	Soluble part of th
10	38	88.4	315	AAU75499	Human fusion prote

1	43	100.0	7	AAY32791	Proteolytic cleavage
2	38	88.4	33	AAR2156	Plasmid pRSR8-3 en
3	38	88.4	246	AAR56389	Human growth hormo
4	38	88.4	249	AAR06367	Hormone binding re
5	38	88.4	269	AAR05045	Soluble human grow
6	38	88.4	269	AAR10426	Human somatogenic
7	38	88.4	269	AAY31767	Human soluble grow
8	38	88.4	269	AAYW82802	Human soluble soma
9	38	88.4	269	AAY78429	Soluble part of th
10	38	88.4	315	AAU75499	Human fusion prote

1	43	100.0	7	AAY32791	Proteolytic cleavage
2	38	88.4	33	AAR2156	Plasmid pRSR8-3 en
3	38	88.4	246	AAR56389	Human growth hormo
4	38	88.4	249	AAR06367	Hormone binding re
5	38	88.4	269	AAR05045	Soluble human grow
6	38	88.4	269	AAR10426	Human somatogenic
7	38	88.4	269	AAY31767	Human soluble grow
8	38	88.4	269	AAYW82802	Human soluble soma
9	38	88.4	269	AAY78429	Soluble part of th
10	38	88.4	315	AAU75499	Human fusion prote

1	43	100.0	7	AAY32791	Proteolytic cleavage
2	38	88.4	33	AAR2156	Plasmid pRSR8-3 en
3	38	88.4	246	AAR56389	Human growth hormo
4	38	88.4	249	AAR06367	Hormone binding re
5	38	88.4	269	AAR05045	Soluble human grow
6	38	88.4	269	AAR10426	Human somatogenic
7	38	88.4	269	AAY31767	Human soluble grow
8	38	88.4	269	AAYW82802	Human soluble soma
9	38	88.4	269	AAY78429	Soluble part of th
10	38	88.4	315	AAU75499	Human fusion prote

1	43	100.0	7	AAY32791	Proteolytic cleavage
2	38	88.4	33	AAR2156	Plasmid pRSR8-3 en
3	38	88.4	246	AAR56389	Human growth hormo
4	38	88.4	249	AAR06367	Hormone binding re
5	38	88.4	269	AAR05045	Soluble human grow
6	38	88.4	269	AAR10426	Human somatogenic
7	38	88.4	269	AAY31767	Human soluble grow
8	38	88.4	269	AAYW82802	Human soluble soma
9	38	88.4	269	AAY78429	Soluble part of th
10	38	88.4	315	AAU75499	Human fusion prote

1	43	100.0	7	AAY32791	Proteolytic cleavage
2	38	88.4	33	AAR2156	Plasmid pRSR8-3 en
3	38	88.4	246	AAR56389	Human growth hormo
4	38	88.4	249	AAR06367	Hormone binding re
5	38	88.4	269	AAR05045	Soluble human grow
6	38	88.4	269	AAR10426	Human somatogenic
7	38	88.4	269	AAY31767	Human soluble grow
8	38	88.4	269	AAYW82802	Human soluble soma
9	38	88.4	269	AAY78429	Soluble part of th
10	38	88.4	315	AAU75499	Human fusion prote

1	43	100.0	7	AAY32791	Proteolytic cleavage
2	38	88.4	33	AAR2156	Plasmid pRSR8-3 en
3	38	88.4	246	AAR56389	Human growth hormo
4	38	88.4	249	AAR06367	Hormone binding re
5	38	88.4	269	AAR05045	Soluble human grow
6	38	88.4	269	AAR10426	Human somatogenic
7	38	88.4	269	AAY31767	Human soluble grow
8	38	88.4	269	AAYW82802	Human soluble soma
9	38	88.4	269	AAY78429	Soluble part of th
10	38	88.4	315	AAU75499	Human fusion prote

1	43	100.0	7	AAY32791	Proteolytic cleavage
2	38	88.4	33	AAR2156	Plasmid pRSR8-3 en
3	38	88.4	246	AAR56389	Human growth hormo
4	38	88.4	249	AAR06367	Hormone binding re
5	38	88.4	269	AAR05045	Soluble human grow
6	38	88.4	269	AAR10426	Human somatogenic
7	38	88.4			

PS	Claim 15; Page 27; 36pp; English.	Matches	6;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
XX	This sequence is a cleavage signal site. This site is used in a method for controlling the availability and signal transduction capability of a cell surface receptor by administering an inhibitor that is capable of inhibiting proteolytic cleavage of the receptor. Inhibition of this proteolytic cleavage results in the receptors being present on the surface for longer and therefore signalling for longer to the interior of the cell. This increases the sensitivity of cells to any hormones which might be present. The inhibitor is derived from or competes with an amino acid sequence around this proteolytic cleavage signal. The inhibitor may be used to treat muscle wasting, associated with disorders such as renal tubular defects, uremia, diabetes, Cushing's syndrome, cachexias, eating disorders, AIDS, after stress and during neuromuscular disease.	Qy	1	CEEDFY	6						
CC	Sequence 7 AA;	Db	28	CEEDFY	33						
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 43; DB 20; Length 7;										
CC	Best Local Similarity 100.0%; Pred. No. 7 8e+05;										
CC	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
CC	OS Homo sapiens.										
CC	Sequence 7 AA;										
CC	Query Match 100.0%; Score 4										

PF 30-JAN-1989; 89JP-0020182.  
 XX  
 PR 30-JAN-1989; 89JP-0020182.  
 XX  
 PA (TANP-) TANPAKU KOGAKU KENK.  
 XX  
 DR WPI: 1990-285858/38.  
 DR N-PSDB; AAQ05968.  
 XX

PT Synthetic gene - used for coding the amino acid sequence in the hormone-combining region of human growth hormone receptor.  
 PT  
 XX Disclosure: Fig 2; 11pp; Japanese.

CC The sequence is the same as that of the natural receptor but is encoded by a synthetic gene that has a slightly altered nucleotide sequence incorporating at least 2 new restriction sites and removing a direct repeat and a palindromic sequence.  
 CC  
 Sequence 249 AA;

Query Match 88.4%; Score 38; DB 11; Length 249;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEEDFY 6  
 Db 241 CEEDFY 246

RESULT 5  
 ID AAR05045 standard; protein: 269 AA.  
 XX  
 AC AAR05045;  
 XX  
 DT 03-OCT-1990 (first entry)  
 XX  
 DE Soluble human growth hormone receptor.  
 XX  
 KW Plasmid pJ1446; human growth hormone; somatogenic receptor; mutagenesis;  
 KW substitutions; active domain; hormone variants.  
 XX  
 OS Synthetic.  
 XX  
 WO9004788-A.  
 XX  
 PD 03-MAY-1990.  
 XX  
 PF 30-OCT-1989; 89WO-US04778.  
 XX  
 PR 29-OCT-1988; 88US-0264611; WO-U004778.  
 XX  
 PA (GETH ) GENENTECH INC.  
 PI Wells JA, Cunningham BC;  
 XX  
 DR WPI: 1990-164120/21.  
 DR N-PSDB; AAQ04672.  
 XX

PT Identifying active domains and amino acid(s) in polypeptide(s) - by specific substitutions, then testing modified products for interaction with target, and new polypeptide, esp. hormone etc.  
 PT  
 PT  
 PT  
 PT  
 PS Disclosure: ; 3pp; English.  
 XX  
 CC The soluble human growth hormone receptor shGHR was subcloned into pBO475 to form pJ1446. E.coli W3110, deGP was transformed with pJ1446 and grown in low phosphate media in a fermentor at 30 degrees C. This 246 amino acid hGHR is produced.  
 CC See also AAQ04671 and AAQ04672.  
 XX  
 Sequence 269 AA;

Query Match 88.4%; Score 38; DB 18; Length 269;  
 Best Local Similarity 100.0%; Pred. No. 16;

Qy 1 CEEDFY 6  
 Db 264 CEEDFY 269

RESULT 6  
 ID AAW10426 standard; protein: 269 AA.  
 XX  
 AC AAW10426;  
 AC 11-AUG-1997 (first entry)  
 XX  
 DE Human somatogenic receptor extracellular domain.  
 XX  
 KW Active site; active domain; growth hormone; somatogenic receptor; mutagenesis.  
 KW  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Peptide  
 FT Peptide  
 /label= Sig-peptide  
 24..269  
 FT Protein  
 FT Mat\_protein  
 /note= "ShGHR (aa1-246)"

Query Match 88.4%; Score 38; DB 11; Length 249;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEEDFY 6  
 Db 241 CEEDFY 246

RESULT 5  
 ID AAR05045 standard; protein: 269 AA.  
 XX  
 AC AAR05045;  
 XX  
 DT 03-OCT-1990 (first entry)  
 XX  
 DE Soluble human growth hormone receptor.  
 XX  
 KW Plasmid pJ1446; human growth hormone; somatogenic receptor; mutagenesis;  
 KW substitutions; active domain; hormone variants.  
 XX  
 OS Synthetic.  
 XX  
 WO9004788-A.  
 XX  
 PD 03-MAY-1990.  
 XX  
 PF 30-OCT-1989; 89WO-US04778.  
 XX  
 PR 29-OCT-1988; 88US-0264611; WO-U004778.  
 XX  
 PA (GETH ) GENENTECH INC.  
 PI Wells JA, Cunningham BC;  
 XX  
 DR WPI: 1990-164120/21.  
 DR N-PSDB; AAQ04672.  
 XX

PT Identifying active domains and amino acid(s) in polypeptide(s) - by specific substitutions, then testing modified products for interaction with target, and new polypeptide, esp. hormone etc.  
 PT  
 PT  
 PT  
 PS Example 3; Fig 12A-C; 86pp; English.  
 XX  
 XX  
 CC A polypeptide (AAW10426) comprises the soluble extracellular domain of human liver growth hormone receptor (somatogenic receptor, shGHR). It was expressed in Escherichia coli transformants using a vector derived from pJ1446 (AT47449). shGHR can be used in a novel method designed for the identification of polypeptide active domains. For human growth hormone (hGH) (see also AAW10425), this involves substituting selected amino acid segments of hGH with analogous segments from analogue polypeptides (human placenta, CC analog, human prolactin and pig growth hormone) and examining the CC effect of the substn. on binding affinity to shGHR. Once active CC site regions are determined, active site amino residues (see also CC AAW10427-62) can be similarly identified.  
 XX  
 SQ \* Sequence 269 AA;

Query Match 88.4%; Score 38; DB 18; Length 269;  
 Best Local Similarity 100.0%; Pred. No. 16;

Matches	6;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;	QY	1 CEEDEFY 6 
QY	1	CEEDEFY 6 								Db	264 CEEDEFY 269
Db	264	CEEDEFY 269									
<b>RESULT 7</b>											
ID	AY31767	AY31767 standard; Protein; 269 AA.									
XX											
AC	AY31767;										
XX											
DT	06-DEC-1999	(first entry)									
XX											
DE	Human soluble growth hormone receptor.										
KW	Growth hormone receptor; somatogenic receptor; human; pJ1446;										
KW	variant; protein engineering.										
XX											
OS	Homo sapiens.										
XX											
Key		Location/Qualifiers									
Peptide		1..23									
ID		/note= "Signal Peptide"									
Protein											
FT											
FT		/note= "mature protein"									
XX											
PN	US5955346-A.										
XX											
PD	21-SEP-1999.										
XX											
PF	07-JUN-1995;	95US-0476999.									
XX											
PR	02-FEB-1994;	94US-0190723.									
PR	26-OCT-1989;	89US-0428062.									
PR	27-APR-1992;	92US-0975204.									
PR	13-OCT-1992;	92US-0875204.									
PR	13-OCT-1992;	92US-096007.									
PR	28-OCT-1988;	89US-0264611.									
XX											
PA	(GETH ) GENENTECH INC.										
XX											
PI	Cunningham BC, Wells JA;										
XX											
DR	WPI: 1999-560495/47.										
XX											
N-PSDB	AAV87977.										
PT	Isolated nucleic acids encoding variants of human prolactin and placental lactogen useful for identifying active domains within those proteins -										
PT											
PT											
Example 2; Fig 12A-C; 86pp; English.											
CC	This sequence represents a human soluble growth hormone receptor (sGHR) encoded by plasmid pJ1446 (see AAI87977). sGHR was expressed in E. coli and was used in binding assays of human growth hormone variants. The invention provides a method for the systematic analysis of the structure and function of polypeptides by identifying active domains which influence the activity of the polypeptide with a target substance, and a method for identifying the active amino acid residues within the active domain of a polypeptide. It also provides polypeptide variants comprising segment-substituted and residue-substituted growth hormones, prolactins (see AAY31764) and placental lactogens (see AY31765).										
CC	Identifying receptor specific variants.										
CC	Sequence 269 AA;										
CC	Best Local Similarity 100.0%; Pred. No. 16; Conservative 6; Mismatches 0; Indels 0; Gaps 0;										
CC	Query Match 88.4%; Score 38; DB 20; Length 269;										
CC	Matched Regions 1-269.										

XX SQ Sequence 269 AA;  
 Query Match 88.4%; Score 38; DB 20; Length 269;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEEDFY 6  
 Db 264 CEEDFY 269

RESULT 9  
 AAY78429 standard; Protein: 269 AA.  
 ID AAY78429  
 AC AAU75499;  
 XX DE Soluble part of the somatogenic receptor encoded by plasmid pJ1446.  
 XX KW Human growth hormone; hGH; prolactin; placental lactogen;  
 KW modification; mutagenesis.  
 XX OS Homo sapiens.  
 OS Synthetic.  
 XX PN US6013478-A.  
 XX PD 11-JAN-2000.  
 XX PF 24-JUN-1998;  
 XX 98US-01040436.  
 PR 26-OCT-1989;  
 PR 99US-0428066.  
 PR 27-APR-1992;  
 PR 92US-0875204.  
 PR 13-OCT-1992;  
 PR 92US-0960227.  
 PR 02-FEB-1994;  
 PR 94US-0190223.  
 PR 06-JUN-1995;  
 PR 95US-0483039.  
 PR 30-JUN-1997;  
 PR 97US-0903398.  
 PR 28-OCT-1988;  
 PR 88US-0264611.  
 XX PA (GETH ) GENENTECH INC.  
 XX PI Wells JA, Cunningham BC;  
 XX DR WPI; 2000-159873/14.  
 XX N-PSDB; AAZ88448.  
 XX PT Recombinant production of variant polypeptides, e.g. growth hormone variants with altered receptor specificity, using cells transformed with DNA selected by scanning mutagenesis in at least one peptide domain.  
 XX PT Example 3; Fig 12; 83pp; English.  
 XX PS The present invention describes the production of a polypeptide variant (I) comprising segment substituted and residue substituted growth hormone, prolactin or placental lactogens. The method is particularly used to produce variants of growth hormone (GH), prolactin or placental lactogen, but may also be applied to receptors, interferons, and colony-stimulating factors. A particular application is the production of human GH variants with altered (decreased or increased) binding interaction with the somatogenic receptor, i.e. compounds useful as human GH (ant)agonists and which may have higher potency for stimulating other human GH receptors, and as standards or tracers in immunoassays for human GH. This method of DNA selection identifies the biologically active residues in active domains, including those critical for interaction with different targets. The present sequence represents the soluble region of the somatogenic receptor from liver encoded by a plasmid, which is used in the exemplification of the present invention.

XX SQ Sequence 269 AA;

Query Match 88.4%; Score 38; DB 21; Length 269;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEEDFY 6  
 Db 264 CEEDFY 269

RESULT 10  
 AAU75499  
 ID AAU75499 standard; Protein: 315 AA.  
 AC AAU75499;  
 XX DE Human fusion protein Chi 1A2.  
 XX KW Human; GHS1-23; human growth hormone; GH; GHR; growth hormone; receptor; GHstopGHR SD100; GHlinkGHR; GHlinkGHRflec; pTrcRNSsacI; TrcHindra; Chi 1A2 chimaera; acromegaly; gigantism; growth hormone deficiency; Turner's syndrome; renal failure; osteoporosis; diabetes mellitus; cancer; obesity; insulin resistance; hyperlipidaemia; hypertension; anaemia; autoimmune disease; infectious disease; inflammatory disorder; rheumatoid arthritis; interleukin-6 chimaera; IL-6.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 XX Misc-difference 14  
 FT /note= "Encoded by ATG"  
 FT Misc-difference 312  
 FT /label= Unknown  
 FT Misc-difference 313  
 FT /note= "Encoded by TGA, in-frame stop codon"  
 FT Misc-difference 313  
 FT /label= Unknown  
 FT Misc-difference 313  
 FT /note= "Encoded by TAA, in-frame stop codon"  
 XX PN WO200196565-A2.  
 XX PD 20-DEC-2001.  
 XX PP 18-JUN-2001; 2001WO GB02645.  
 XX PR 16-JUN-2000; 2000GB-0014765.  
 XX PR 10-MAR-2001; 2001GB-0005669.  
 XX PR 16-MAR-2001; 2001GB-0006487.  
 XX PA (ASTE-) ASTERION LTD.  
 XX PI Ross R, Artymuk P, Sayers J;  
 XX DR WPI; 2002-130734/17.  
 XX N-PSDB; ABR14547.  
 XX New binding agent useful in producing a medicament for treating e.g. cancer, obesity, acromegaly or gigantism, comprises a first part that binds to a ligand binding domain of a receptor and a second part having a receptor binding domain.  
 XX PS Claim 49; Fig 22; 79pp; English.  
 XX CC The invention relates to a binding agent comprising a first part capable of binding ligand binding domain of a receptor linked to a second part comprising a receptor binding domain, where the binding agent modulates the activity of the receptor. Also included are a nucleic acid molecule having a sequence, which encodes a binding agent comprising sequences given in the specification comprising the sequences of the full length GHstopGHR SD100 construct, GHlinkGHR construct(GH, growth hormone, GHR, growth hormone receptor), GHlinkGHRflec construct, 1157 base pair

CC PCR fragment GHlinkGHR generated by nucleotides pTrCRN5Saci and  
 CC Trchindrev, or the nucleotide sequence of the Chi 1A2 chimaera,  
 CC sequences binding to the nucleic acids or degenerate sequences  
 CC representing them (which have receptor antagonising activity),  
 CC their encoded polypeptides, a vector comprising the nucleic acids and a  
 CC cell transformed/transfected with the nucleic acid or vector.  
 CC the binding agent is useful for manufacturing a medicament for  
 CC the treatment of acromegaly, gigantism, growth hormone (GH) deficiency,  
 CC Turner's syndrome, renal failure, osteoporosis, diabetes mellitus,  
 CC cancer, obesity, insulin resistance, hyperlipidaemia, hypertension,  
 CC anaemia, autoimmune and infectious diseases, and inflammatory disorders  
 CC including rheumatoid arthritis (interleukin (IL)-6 chimaera).  
 CC The present sequence represents the fusion protein agent of the invention  
 CC being the Chi 1A2 construct (growth hormone/growth hormone  
 CC receptor not linked by a synthetic peptide linker).

XX Sequence 315 AA;  
 XX Query Match 88.4%; Score 38; DB 23; Length 315;  
 XX Best Local Similarity 100.0%; Pred. No. 18;  
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX AC 1 CEBDFY 6  
 DB 306 CEBDFY 311

RESULT 11  
 AAU75496 Standard; Protein: 340 AA.  
 XX AC AAU75496;  
 XX DT 08-MAY-2002 (first entry)

DE Human fusion protein GHlinkGHR.  
 XX Human; GH51-23; Human growth hormone; GH; GHR;  
 KW growth hormone; receptor; GHstopGHR SD100; GHlinkGHRfilec;  
 KW pTrCRN5Saci; Trchindrev; Chi 1A2 chimaera; acromegaly; gigantism;  
 KW growth hormone deficiency; Turner's syndrome; renal failure;  
 KW osteoporosis; diabetes mellitus; cancer; obesity; insulin resistance;  
 KW hyperlipidaemia; hypertension; anaemia; autoimmune disease;  
 KW infectious disease; inflammatory disorder; rheumatoid arthritis;  
 KW interleukin-6 chimaera; IL-6.

XX OS Homo sapiens.  
 OS Synthetic.

XX FH Key

Misc-difference 14

Location/Qualifiers

Misc-difference 337

/note= "Encoded by ATG"

Misc-difference 338

/label= Unknown

/note= "Encoded by TAA, in-frame stop codon"

FT XW WO200106565-A2.

PN 20-DEC-2001.

XX 18-JUN-2001; 2001WO-GB02645.

PD 20-DEC-2001.

XX 16-JUN-2000; 2000GB-0014765.

PR 10-MAR-2001; 2001GB-0005969.

PR 16-MAR-2001; 2001GB-0006487.

XX (ASTE-) ASTERION LTD.

XX Ross R. Artymiuik P., Sayers J;

XX Hammonds RG, Leung DW, Martin DW, Spencer SA, Wood WI;

WPI; 2002-130734/17.  
 DR N-PSDB; ABK14531.

XX PT New binding agent useful in producing a medicament for treating e.g. cancer, obesity, acromegaly or gigantism, comprises a first part that binds to a ligand binding domain of a receptor and a second part having a receptor binding domain -

XX Disclosure; Fig 6: 79pp; English.

XX The invention relates to a binding agent comprising a first part capable of binding a ligand binding domain of a receptor linked to a second part comprising a receptor binding domain, where the binding agent modulates the activity of the receptor. Also included are a nucleic acid molecule having a sequence, which encodes a binding agent comprising sequences given in the specification comprising the sequences of the full length

CC GHstopGHR SD100 construct, GHlinkGHRfilec construct, 1157 base pair

CC PCR fragment GHlinkGHR generated by nucleotides pTrCRN5Saci and

CC Trchindrev, or the nucleotide sequence of the Chi 1A2 chimaera, sequences binding to the nucleic acids or degenerate sequences

CC representing them (which have receptor antagonising activity), their encoded polypeptides, a vector comprising the nucleic acids and a

CC cell transformed/transfected with the nucleic acid or vector.

CC The binding agent is useful for manufacturing a medicament for the treatment of acromegaly, gigantism, growth hormone (GH) deficiency, Turner's syndrome, renal failure, osteoporosis, diabetes mellitus, cancer, obesity, insulin resistance, hyperlipidaemia, hypertension, anaemia, autoimmune and infectious diseases, and inflammatory disorders including rheumatoid arthritis (interleukin (IL)-6 chimaera).

CC The present sequence represents the fusion protein agent of the invention

CC being the GHlinkGHR construct (growth hormone/growth hormone receptor linked by a synthetic peptide linker).

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;

XX Query Match 88.4%; Score 38; DB 23; Length 340;

XX Best Local Similarity 100.0%; Pred. No. 20;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AC 1 CBEDFY 6

DB 331 CEEEDFY 336

XX Sequence 340 AA;</



CC efficacy of growth hormone in vivo.  
 XX Sequence 638 AA;  
 Query Match Score 38; DB 19; Length 638;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 CEEDFY 6  
 Db 259 CEEEDFY 264

RESULT 15  
 ABB11437 standard; peptide: 648 AA.  
 ID ABB11437  
 XX AC ABB11437;  
 XX DT 11-JAN-2002 (first entry)  
 XX DE Human growth hormone receptor homologue, SEQ ID NO:1807.  
 Human: cytokine; cell proliferation; cell differentiation; growth factor;  
 haemopoiesis regulation; tissue growth; immunomodulator; activin;  
 KW inhibin; chemotaxis; chemokines; thrombolytic; oncogenesis;  
 KW proliferation; metastasis; cancer; tumour; haemopoietic disorder;  
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
 KW chronic inflammatory condition; proliferative retinopathy;  
 KW atherosclerosis; coronary heart disease; arterial ischaemia;  
 KW bone disorder; osteoporosis; vascular growth disorder;  
 KW tissue regeneration; wound healing; infection; immune disorder;  
 KW cell culture; drug screening; gene therapy; antiinflammatory;  
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;  
 KW cytostatic; osteopathic; vasotropic; cardiant; viricide; antibacterial;  
 KW antifungal; vulnerary; antilulcer.  
 XX Homo sapiens.  
 OS WO200157188-A2.  
 PN XX  
 PD XX  
 PF 09-AUG-2001.  
 XX  
 PR 05-FEB-2001; 2001WO-US03800.  
 XX  
 PR 03-FEB-2000; 2000US-0496914.  
 XX  
 PR 27-APR-2000; 2000US-0560875.  
 XX PA (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Drmanac RT;  
 PS WPI; 2001-457740/49.  
 DR N-PSDB; ABA08681.

XX Human Proteins and DNA encoding sequences useful for preventing,  
 PT treating or ameliorating a medical condition in a mammalian subject  
 PR e.g. arthritis and cancer -  
 XX  
 PS Claim 20: Page 189-190; 1963pp; English.  
 XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
 CC invention also relates to vectors and recombinant host cells comprising a  
 CC nucleotide of the invention, methods of producing the novel polypeptides,  
 CC antibodies against the polypeptides, methods of detecting the nucleotides  
 CC or polypeptides in a sample, and methods of identifying compounds which  
 CC bind to polypeptides of the invention. Although novel, many of the  
 CC polypeptides of the invention have homology to known proteins, thereby  
 CC giving an insight into their probable biological activities, and hence  
 CC potential therapeutic applications. The polypeptides of the invention may  
 CC have various activities, including cytokine, cell proliferation or cell  
 CC differentiation activities; stem cell growth factor activity;

CC haemopoiesis regulatory activity; tissue growth activity;  
 CC immunomodulatory activity; activin- or inhibitin-related activities;  
 CC chemotactic or chemoattractive activities; haemostatic, thrombotic or  
 CC thrombolytic activities; receptor or ligand activities; or may be  
 CC involved in oncogenesis, cancer cell proliferation or metastasis.  
 CC Depending on their biological activities, polypeptides and nucleotides of  
 CC the invention are useful for preventing, treating or ameliorating medical  
 CC conditions, e.g., by protein or gene therapy. Such conditions include  
 CC cancers, haemopoietic disorders (e.g., myeloid or lymphoid cell  
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
 CC vascular growth. Polypeptides involved with tissue regeneration and  
 CC repair (or nucleic acids encoding them) may be used to promote wound  
 CC healing (e.g., of burns, incisions and ulcers), while those with  
 CC immunomodulatory activities may be used in the treatment of viral,  
 CC bacterial and fungal infections in addition to immune disorders.  
 CC Polypeptides with growth factor activity may be used in cell cultures to  
 CC promote cell growth. For example, such polypeptides may be used to  
 CC manipulate stem cells in culture to give rise to neuroepithelial cells  
 CC that can be used to augment or replace cells damaged by illness.  
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
 CC may also be used in the diagnosis of the above conditions, and in drug  
 CC screening techniques. The present sequence represents a novel human  
 CC polypeptide of the invention.  
 XX

SQ Sequence 648 AA;

Query	Match	Score 38;
Best Local	Similarity 100.0%;	Pred. No. 38;
Matches 6;	Conservative 0;	Mismatches 0;
QY	1 CEEDFY 6	Db 269 CEEEDFY 274

Search completed: February 12, 2003, 11:44:39  
 Job time : 30 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: February 12, 2003, 11:43:40 ; Search time 10.7333 Seconds  
(without alignments)  
19.189 Million cell updates/sec

Title: US-09-660-302c-7

Perfect score: 43  
Sequence: 1 CEDDFYR 7

Scoring table: BLOSUM62

Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters:

262574

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:  
1: /cgn2\_6/podata/1/1aa/5A\_COMB.pep:  
2: /cgn2\_6/podata/1/1aa/5B\_COMB.pep:  
3: /cgn2\_6/podata/1/1aa/6A\_COMB.pep:  
4: /cgn2\_6/podata/1/1aa/6B\_COMB.pep:  
5: /cgn2\_6/podata/1/1aa/pcrUS\_COMB.pep:  
6: /cgn2\_6/podata/1/1aa/backfiles1.pep:  
\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	38	88.4	620	4	US-09-000-145-1		Sequence 1, Appli
2	36	83.7	84	2	US-08-465-380-59		Sequence 59, Appli
3	36	83.7	84	2	US-08-486-397-59		Sequence 59, Appli
4	36	83.7	84	2	US-08-486-399-59		Sequence 59, Appli
5	36	83.7	84	2	US-08-461-965-59		Sequence 59, Appli
6	36	83.7	84	2	US-08-534-601-59		Sequence 59, Appli
7	36	83.7	84	3	US-09-249-471-59		Sequence 59, Appli
8	36	83.7	84	3	US-09-249-472-59		Sequence 59, Appli
9	36	83.7	84	3	US-09-249-471-59		Sequence 59, Appli
10	36	83.7	84	3	US-08-809-455-59		Sequence 59, Appli
11	36	83.7	84	3	US-09-249-461-59		Sequence 59, Appli
12	36	83.7	84	3	US-09-249-448-59		Sequence 59, Appli
13	36	83.7	91	2	US-08-465-380-128		Sequence 128, Appli
14	36	83.7	91	2	US-08-480-478-50		Sequence 50, Appli
15	36	83.7	91	2	US-08-486-397-128		Sequence 128, Appli
16	36	83.7	91	2	US-08-486-399-128		Sequence 128, Appli
17	36	83.7	91	2	US-08-461-965-128		Sequence 128, Appli
18	36	83.7	91	2	US-08-526-110A-50		Sequence 50, Appli
19	36	83.7	91	3	US-08-534-641-128		Sequence 128, Appli
20	36	83.7	91	3	US-09-249-471-128		Sequence 128, Appli
21	36	83.7	91	3	US-09-249-472-128		Sequence 128, Appli
22	36	83.7	91	3	US-09-249-451-128		Sequence 128, Appli
23	36	83.7	91	3	US-08-809-455-128		Sequence 128, Appli
24	36	83.7	91	3	US-09-249-441-128		Sequence 128, Appli
25	36	83.7	91	3	US-09-249-448-128		Sequence 128, Appli
26	34	79.1	723	6	5200183-4		Patent No. 5200183
27	33	76.7	75	2	US-08-465-380-6		Sequence 6, Appli

## ALIGNMENTS

RESULT 1	US-09-000-145-1						
	; Sequence 1, Application US/09000145						
	; Patent No. 6169172						
	; GENERAL INFORMATION:						
	; APPLICANT: DEVAUCHELLE, Gerard						
	; APPLICANT: GARNIER, Laurence						
	; APPLICANT: CAHOREAU, Claire						
	; APPLICANT: CERUTTI, Martine						
	; TITLE OF INVENTION: USE OF A PROLACTIN RECEPTOR OR GROWTH HORMONE RECEPTOR						
	; FILE REFERENCE: 0384-0047-0XPCT						
	; CURRENT APPLICATION NUMBER: US/09/000,145						
	; CURRENT FILING DATE: 1998-03-16						
	; EARLIER APPLICATION NUMBER: PCT/FR96/01237						
	; EARLIER FILING DATE: 1996-08-02						
	; EARLIER APPLICATION NUMBER: FR 95/09420						
	; EARLIER FILING DATE: 1995-08-02						
	; NUMBER OF SEQ ID NOS: 6						
	; SOFTWARE: PatentIn Ver. 2.0						
	; SEQ ID NO: 1						
	; LENGTH: 620						
	; TYPE: PT						
	; ORGANISM: Homo sapiens						
	; US-09-000-145-1						
	Query Match	88.4%	Score 38;	DB 4;	Length 620;		
	Best Local Similarity	100.0%	Pred. No. 15;				
	Matches 6; Conservative, 0; Mismatches 0; Indels 0; Gaps 0;						

STREET: 633 West Fifth Street  
 STREET: Suite 4700  
 CITY: Los Angeles  
 STATE: California  
 ZIP: U.S.A.  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: storage  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: IBM P.C. DOS 5.0  
 SOFTWARE: Word Perfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/465,380  
 FILING DATE: June 5, 1995  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/326,110  
 FILING DATE: October 18, 1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BIGGS, SUZANNE L.  
 REGISTRATION NUMBER: 30,158  
 REFERENCE/DOCKET NUMBER: 213/769  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (213) 489-1600  
 TELEFAX: (213) 955-0440  
 TELEX: 67-3510  
 INFORMATION FOR SEQ ID NO: 59:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 84 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 ORIGINAL SOURCE: peptide  
 ORGANISM: Anyclostoma caninum  
 US-08-465-380-59

Query Match 83.7%; Score 36; DB 2; Length 84;  
 Best Local Similarity 85.7%; Pred. No. 5;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CBEDFYR 7  
 Db 52 CBEGFYR 58

RESULT 3  
 US-08-465-397-59  
 Sequence 59 Application US/08486397  
 PATENT NO. 5866542  
 GENERAL INFORMATION:  
 APPLICANT: George P. Vlasuk, Patric H. Stanssens,  
 Joris H.L. Mensens, Marc J. Lauwers, Yves R. Laroch, Laurent S. Jespers,  
 Yannick G.J. Gansmans, Matthew Moyle,  
 APPLICANT: Peter W. Bergum  
 TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT  
 NUMBER OF SEQUENCES: 357  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Lyon & Lyon  
 STREET: 633 West Fifth Street  
 CITY: Los Angeles  
 STATE: California  
 ZIP: 90071  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: IBM P.C. DOS 5.0  
 SOFTWARE: Word Perfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/486,399  
 FILING DATE: June 5, 1995  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/326,110  
 FILING DATE: October 18, 1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BIGGS, SUZANNE L.  
 REGISTRATION NUMBER: 30,158  
 REFERENCE/DOCKET NUMBER: 213/770  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600  
 TELEFAX: (213) 955-0440  
 TELEX: 67-3510  
 INFORMATION FOR SEQ ID NO: 59:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 84 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: Peptide  
 ORIGINAL SOURCE: Ancylostoma caninum

US-08-486-399-59

Query Match 83.7%; Score 36; DB 2; Length 84;  
 Best Local Similarity 85.7%; Pred. No. 5; Mismatches 0; Gaps 0;

Qy 1 CEEDEFYR 7	Db 52 CEEGFYR 58
-----------------	------------------

RESULT 5

US-08-461-965-59

Query Match 83.7%; Score 36; DB 2; Length 84;  
 Best Local Similarity 85.7%; Pred. No. 5; Mismatches 0; Gaps 0;

Qy 1 CEEDEFYR 7	Db 52 CEEGFYR 58
-----------------	------------------

RESULT 6

US-08-634-641-59

Query Match 83.7%; Score 36; DB 2; Length 84;  
 Best Local Similarity 85.7%; Pred. No. 5; Mismatches 0; Gaps 0;

Qy 1 CEEDEFYR 7	Db 52 CEEGFYR 58
-----------------	------------------

GENERAL INFORMATION:

APPLICANT: Vlaasuk, George P. Vlaasuk  
 ADDRESS: Stanssens, Patrick Eric Hugo  
 Menseens, Joris Hilda Lieven  
 Lauwersys, Marc Josef  
 Laroche, Yves Rene  
 Jespers, Laurent Stephane  
 Gansmans, Yannick Georges Jozef  
 Bergum, Peter W.  
 Moyle, Matthew  
 Bergum, Peter W.

TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT

NUMBER OF SEQUENCES: 356

CORRESPONDENCE ADDRESS:

ADDRESS: Lyon & Lyon  
 STREET: 633 West Fifth Street  
 SUITE: Suite 4700  
 CITY: Los Angeles  
 STATE: California  
 COUNTRY: U.S.A.  
 ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: IBM P.C. DOS 5.0  
 SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US-08/634,641  
 FILING DATE: April 19, 1996  
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/13231  
 FILING DATE: October 17, 1995  
 APPLICATION NUMBER: 08/486,399  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/486,397  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/465,380  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/461,965  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/326,110  
 FILING DATE: October 18, 1994  
 ATTORNEY/AGENT INFORMATION:

NAME: BIGGS, SUSANNE L.  
 REGISTRATION NUMBER: 30,158  
 REFERENCE/DOCKET NUMBER: 210/243  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600  
 TELEFAX: (213) 955-0440  
 TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 59:

SEQUENCE CHARACTERISTICS:

LENGTH: 84 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 ORIGINAL SOURCE: Ancylostoma caninum

ORGANISM: Ancylostoma caninum

US-08-634-641-59  
 Query Match 83.7%; Score 36; DB 2; Length 84  
 Best Local Similarity 85.7%; Pred. No. 5;  
 Matches 6; Conservative 0; Mismatches 1;  
 Indels 0; Gaps 0;

Qy 1 CEEDEFYR 7  
 Db 52 CEEGFYR 58

---

RESULT 7  
 US-09-249-471-59  
 ; Sequence 59, Application US/09249471  
 ; Patent No. 604041  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Vlasuk, George Phillip  
 ; APPLICANT: Stanssens, Patrick Eric Hugo  
 ; APPLICANT: Messens, Joris Hilda Lieven  
 ; APPLICANT: Lauwers, Marc Josef  
 ; APPLICANT: Laroche, Yves Rene  
 ; APPLICANT: Jespers, Laurent Stephane  
 ; APPLICANT: Gansseman, Yannick Georges Jozef  
 ; TITLE OF INVENTION: NEMATODE EXTRACTED SERINE PROTEASE  
 ; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
 ; NUMBER OF SEQUENCES: 356  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Lyon & Lyon  
 ; STREET: 633 West Fifth Street  
 ; STREET: Suite 4700  
 ; CITY: Los Angeles  
 ; STATE: California  
 ; COUNTRY: U.S.A.  
 ; ZIP: 90071  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: IBM P.C. DOS 5.0  
 ; SOFTWARE: Word Perfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/249,471  
 ; FILING DATE:  
 ; APPLICATION NUMBER: 08/809,455  
 ; FILING DATE: April 17, 1997  
 ; APPLICATION NUMBER: PCT/US95/13231  
 ; FILING DATE: October 17, 1995  
 ; APPLICATION NUMBER: 08/486,399  
 ; FILING DATE: June 5, 1995  
 ; APPLICATION NUMBER: 08/486,397  
 ; FILING DATE: June 5, 1995  
 ; APPLICATION NUMBER: 08/465,380  
 ; FILING DATE: June 5, 1995  
 ; APPLICATION NUMBER: 08/461,965  
 ; FILING DATE: June 5, 1995  
 ; APPLICATION NUMBER: 08/461,966  
 ; FILING DATE: June 5, 1995  
 ; APPLICATION NUMBER: 08/461,967  
 ; FILING DATE: October 18, 1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BIGGS, SUZANNE L.  
 ; REGISTRATION NUMBER: 30,158  
 ; REFERENCE/DOCKET NUMBER: 216/270  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (213) 489-1600  
 ; TELEFAX: (213) 955-0440  
 ; TELEX: 67-3510  
 ; INFORMATION FOR SEQ ID NO: 59:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 84 amino acids  
 ; TYPE: amino acid

US-09-249-472-59  
 ; Sequence 59, Application US/09249472  
 ; Patent No. 6040418  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Vlasuk, George Phillip  
 ; APPLICANT: Stanssens, Patrick Eric Hugo  
 ; APPLICANT: Messens, Joris Hilda Lieven  
 ; APPLICANT: Lauwers, Marc Josef  
 ; APPLICANT: Laroche, Yves Rene  
 ; APPLICANT: Jespers, Laurent Stephane  
 ; APPLICANT: Gansseman, Yannick Georges Jozef  
 ; APPLICANT: Moyle, Matthew W.  
 ; APPLICANT: Bergum, Peter W.  
 ; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
 ; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
 ; NUMBER OF SEQUENCES: 356  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Lyon & Lyon  
 ; STREET: 633 West Fifth Street  
 ; STREET: Suite 4700  
 ; CITY: Los Angeles  
 ; STATE: California  
 ; COUNTRY: U.S.A.  
 ; ZIP: 90071  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: IBM P.C. DOS 5.0  
 ; SOFTWARE: Word Perfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/249,472  
 ; FILING DATE:  
 ; APPLICATION NUMBER: 08/809,455  
 ; FILING DATE: April 17, 1997  
 ; APPLICATION NUMBER: PCT/US95/13231  
 ; FILING DATE: October 17, 1995  
 ; APPLICATION NUMBER: 08/486,399  
 ; FILING DATE: June 5, 1995  
 ; APPLICATION NUMBER: 08/465,380  
 ; FILING DATE: June 5, 1995  
 ; APPLICATION NUMBER: 08/461,965  
 ; FILING DATE: June 5, 1995  
 ; APPLICATION NUMBER: 08/461,966  
 ; FILING DATE: June 5, 1995  
 ; APPLICATION NUMBER: 08/461,967  
 ; FILING DATE: June 5, 1995  
 ; APPLICATION NUMBER: 08/465,380  
 ; FILING DATE: June 5, 1995  
 ; APPLICATION NUMBER: 08/461,966  
 ; FILING DATE: June 5, 1995  
 ; APPLICATION NUMBER: 08/326,110  
 ; FILING DATE: October 18, 1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BIGGS, SUZANNE L.  
 ; REGISTRATION NUMBER: 30,158  
 ; REFERENCE/DOCKET NUMBER: 216/270  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (213) 489-1600  
 ; TELEFAX: (213) 955-0440  
 ; TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 59:

SEQUENCE CHARACTERISTICS:

LENGTH: 84 amino acids

TYPE: amino acid

TOPOLGY: linear

MOLECULE TYPE: Peptide

ORIGINAL SOURCE:

ORGANISM: Ancylostoma caninum

US-09-29-472-59

Query Match Score 36; DB 3; Length 84;

Best Local Similarity 85.7%; Pred. No. 5;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEDDFYR 7

RESULT 9

Sequence 59, Application US/09249451

PATENT NO. 6087487

GENERAL INFORMATION:

APPLICANT: Vlasuk, George Phillip

APPLICANT: Stanssens, Patrick Eric Hugo

APPLICANT: Messens, Joris Hilda Lieven

APPLICANT: Lauwereys, Marc Josef

APPLICANT: Larroche, Yves Rene

APPLICANT: Jaspers, Laurent Stephane

APPLICANT: Gansmans, Yannick Georges Jozef

APPLICANT: Moyle, Matthew

APPLICANT: Bergum, Peter W.

TITLE OF INVENTION: NEMATODE-EXTRACED SERINE PROTEASE

TITLE OF INVENTION: INHIBITORS AND ANTICOGULANT

TITLE OF INVENTION: PROTEIN

NUMBER OF SEQUENCES: 356

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C., DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/249,451

FILING DATE: April 17, 1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/13231

FILING DATE: October 17, 1995

APPLICATION NUMBER: 08/486,399

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/486,397

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/465,380

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/461,965

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/326,110

FILING DATE: October 18, 1994

ATTORNEY/AGENT INFORMATION:

NAME: BIGGS, SUZANNE L.

REGISTRATION NUMBER: 30-158

REFERENCE/DOCKET NUMBER: 216/270

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 483-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

SEQUENCE CHARACTERISTICS:

LENGTH: 84 amino acids

TYPE: amino acid

TOPOLGY: linear

MOLECULE TYPE: Peptide

ORIGINAL SOURCE:

ORGANISM: Ancylostoma caninum

US-09-249-451-59

Query Match Score 36; DB 3; Length 84;

Best Local Similarity 85.7%; Pred. No. 5;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEDDFYR 7

RESULT 10

Sequence 59, Application US/08809455

PATENT NO. 6090316

GENERAL INFORMATION:

APPLICANT: Viasuk, George Phillip

APPLICANT: Stanssens, Patrick Eric Hugo

APPLICANT: Messens, Joris Hilda Lieven

APPLICANT: Lauwereys, Marc Josef

APPLICANT: Larroche, Yves Rene

APPLICANT: Jaspers, Laurent Stephane

APPLICANT: Gansmans, Yannick Georges Jozef

APPLICANT: Moyle, Matthew

APPLICANT: Bergum, Peter W.

TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE

TITLE OF INVENTION: INHIBITORS AND ANTICOGULANT

NUMBER OF SEQUENCES: 356

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C., DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/809,455

FILING DATE: April 17, 1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/13231

FILING DATE: October 17, 1995

APPLICATION NUMBER: 08/486,399

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/486,397

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/465,380

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/461,965

FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/326,110

FILING DATE: October 18, 1994

ATTORNEY/AGENT INFORMATION:

NAME: BIGGS, SUZANNE L.

REGISTRATION NUMBER: 30-158

REFERENCE/DOCKET NUMBER: 216/270

NAME: BIGGS, SUZANNE L.

REGISTRATION NUMBER: 30,158  
 REFERENCE/DOCKET NUMBER: 216/270  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (213) 489-1600  
 TELEX: (213) 955-0440  
 TELLEX: 67-3510  
 INFORMATION FOR SEQ ID NO: 59:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 84 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 ORIGINAL SOURCE:  
 ORGANISM: Anyclostoma caninum  
 US-08-809-455-59

Query Match 83.7%; Score 36; DB 3; Length 84;  
 Best Local Similarity 85.7%; Pred. No. 5;  
 Matches 6; Conservative 0; Mismatches 0;  
 Gaps 0;

Qy 1 CEEFDYR 7  
 |||||  
 52 CEEFDYR 58

---

RESULT 11  
 US-09-249-461-59  
 Sequence 59, Application US/09249461  
 Patent No. 6096877  
 GENERAL INFORMATION:  
 APPLICANT: Vlasek, George Phillip  
 APPLICANT: Stanssens, Patrick Eric Hugo  
 APPLICANT: Messens, Joris Hilda Lieven  
 APPLICANT: Lauwers, Marc Joseph  
 APPLICANT: Laroche, Yves Rene  
 APPLICANT: Jespers, Laurent Stephane  
 APPLICANT: Ganssemans, Yannick Georges Jozef  
 APPLICANT: Moyle, Matthew  
 APPLICANT: Bergum, Peter W.  
 TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
 TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT  
 TITLE OF INVENTION: PROTEIN  
 NUMBER OF SEQUENCES: 356  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Lyon & Lyon  
 STREET: 633 West Fifth Street  
 STREET: Suite 4700  
 CITY: Los Angeles  
 STATE: California  
 COUNTRY: U.S.A.  
 ZIP: 90071

COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 MEDIUM TYPE: storage  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: IBM P.C., DOS 5.0  
 SOFTWARE: Word Perfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/249,461  
 FILING DATE:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/809,455  
 FILING DATE: April 17, 1997  
 APPLICATION NUMBER: PCT/US95/13231  
 FILING DATE: October 17, 1995  
 APPLICATION NUMBER: 08/486,399  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/486,397  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/465,380  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/461,965  
 FILING DATE: June 5, 1995

---

APPLICATION NUMBER: 08/326,110  
 FILING DATE: October 18, 1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BIGGS, SUZANNE L.  
 REGISTRATION NUMBER: 30,158  
 REFERENCE/DOCKET NUMBER: 216/270  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (213) 489-1600  
 TELEFAX: (213) 955-0440  
 TELEX: 67-3510  
 INFORMATION FOR SEQ ID NO: 59:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 84 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 ORIGINAL SOURCE:  
 ORGANISM: Anyclostoma caninum  
 US-09-249-461-59

Query Match 83.7%; Score 36; DB 3; Length 84;  
 Best Local Similarity 85.7%; Pred. No. 5;  
 Matches 6; Conservative 0; Mismatches 1;  
 Gaps 0;

Qy 1 CEEFDYR 7  
 |||||  
 Db 52 CEEFDYR 58

---

RESULT 12  
 US-09-249-448-59  
 Sequence 59, Application US/09249448  
 Patent No. 6121435  
 GENERAL INFORMATION:  
 APPLICANT: Vlasek, George Phillip  
 APPLICANT: Stanssens, Patrick Eric Hugo  
 APPLICANT: Messens, Joris Hilda Lieven  
 APPLICANT: Lauwers, Marc Joseph  
 APPLICANT: Laroche, Yves Rene  
 APPLICANT: Jespers, Laurent Stephane  
 APPLICANT: Ganssemans, Yannick Georges Jozef  
 APPLICANT: Moyle, Matthew  
 APPLICANT: Bergum, Peter W.  
 TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE  
 NUMBER OF SEQUENCES: 356  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Lyon & Lyon  
 STREET: 633 West Fifth Street  
 STREET: Suite 4700  
 CITY: Los Angeles  
 STATE: California  
 COUNTRY: U.S.A.  
 CITY: Los Angeles  
 STATE: California  
 COUNTRY: U.S.A.  
 ZIP: 90071

COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 MEDIUM TYPE: storage  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: IBM P.C., DOS 5.0  
 SOFTWARE: Word Perfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/249,448  
 FILING DATE:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/809,455  
 FILING DATE: April 17, 1997  
 APPLICATION NUMBER: PCT/US95/13231  
 FILING DATE: October 17, 1995  
 APPLICATION NUMBER: 08/465,399  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/486,397  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/461,965  
 FILING DATE: June 5, 1995

APPLICATION NUMBER: 08/465,380  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/461,965  
 FILING DATE: June 5, 1995  
 APPLICATION NUMBER: 08/326,110  
 FILING DATE: October 18, 1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BIGGS, SUZANNE L.  
 REGISTRATION NUMBER: 30,158  
 REFERENCE/DOCKET NUMBER: 216/270  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (213) 489-1600  
 TELEX: (213) 955-0440  
 TELEX: 67-3510  
 INFORMATION FOR SEQ ID NO: 59:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 84 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 ORIGINAL SOURCE: ORGANISM: Ancylostoma caninum  
 US-09-249-448-59

Query Match Score 83.7%; DB 3; Length 84;  
 Best Local Similarity 85.7%; Pred. No. 5;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CEEDFYR 7  
 Db 52 CEEGFYR 58

RESULT 13  
 US-08-465-380-128  
 Sequence 128, Application US/08465380  
 Patent No. 5863894  
 GENERAL INFORMATION:  
 APPLICANT: George P. Vlasuk, Patric H. Stanssens,  
 APPLICANT: Joris H.L. Mensens, Marc J. Lauwersys,  
 APPLICANT: Yves R. Laroche, Laurent S. Jespers,  
 APPLICANT: Yannick G.J. Gansens, Matthew Moyle,  
 APPLICANT: Peter W. Bergum  
 TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT  
 TITLE OF INVENTION: PROTEIN  
 NUMBER OF SEQUENCES: 356  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Lyon & Lyon  
 STREET: 633 West Fifth Street  
 STREET: Suite 4700  
 CITY: Los Angeles  
 STATE: California  
 COUNTRY: U.S.A.  
 ZIP: 90071

COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 MEDIUM TYPE: storage  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: IBM P.C. DOS 5.0  
 SOFTWARE: WORD Perfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/465,380  
 FILING DATE: June 5, 1995  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/326,110  
 FILING DATE: October 18, 1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BIGGS, SUZANNE L.  
 REGISTRATION NUMBER: 30,158  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (213) 489-1600

Query Match Score 83.7%; DB 3; Length 91;  
 US-08-465-380-128  
 Sequence 50, Application US/08480478  
 Patent No. 5864009  
 GENERAL INFORMATION:  
 APPLICANT: PATRICK ERIC  
 APPLICANT: HUGO STANSSENS; JORIS HILDA  
 APPLICANT: LIEVEN MESENS; MARC JOZEF  
 APPLICANT: LAUWREYS; YVES RENE LAROCHE;  
 APPLICANT: LAURENT STEPHANE JESPERS; and  
 APPLICANT: YANNICK GEORGES JOZEF  
 APPLICANT: GANSEMANS  
 TITLE OF INVENTION: NEMATODE-EXTRACTED ANTI-  
 TITLE OF INVENTION: COAGULANT PROTEIN  
 NUMBER OF SEQUENCES: 86  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Lyon & Lyon  
 STREET: Suite 4700  
 STREET: 633 West Fifth Street  
 CITY: Los Angeles  
 STATE: California  
 COUNTRY: U.S.A.  
 ZIP: 90071  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 MEDIUM TYPE: storage  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: IBM P.C. DOS 5.0  
 SOFTWARE: FASTSEQ Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/480,478  
 FILING DATE: 06-JUN-1995  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 530  
 CLASSIFICATION: 530  
 FILING DATE: 18-OCTOBER-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BIGGS, SUZANNE L.  
 REGISTRATION NUMBER: 30,158  
 REFERENCE/DOCKET NUMBER: 208/290  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (213) 489-1600  
 TELEFAX: (213) 955-0440  
 TELEX: 67-3510  
 INFORMATION FOR SEQ ID NO: 50:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 91 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 US-08-480-478-50

Query Match Score 83.7%; DB 2; Length 91;

Best Local Similarity 85.7%; Pred. No. 5.4;  
Matches 6; Conservative 0; Mismatches

Qy 1 CEEDEYR 7  
Db 59 CEEGFYR 65

## RESULT 15 US-08-486-397-128

; Sequence 128, Application US/08486397  
; Patent No. 5066542

## GENERAL INFORMATION:

; APPLICANT: George P. Vlasuk, Patrick H. Stanssens,  
; Joris H.L. Mensens, Marc J. Lauverys,  
; Yves R. Laroche, Laurent S. Jespers,  
; Yannick G.J. Gansmans, Matthew Moyle,  
; Peter W. Bergum  
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT  
; NUMBER OF SEQUENCES: 357

## CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; SUITE: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.

## ZIP: 90071

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

## ; COMPUTER: IBM Compatible

; OPERATING SYSTEM: IBM P.C., DOS 5.0  
; SOFTWARE: Word Perfect 5.1

## ; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,397  
; FILING DATE: June 5, 1995  
; CLASSIFICATION: 530

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/326,110  
; FILING DATE: October 18, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BIGGS, SUZANNE L.

; REGISTRATION NUMBER: 30,158  
; REFERENCE/DOCKET NUMBER: 213/269

; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEX: 67-3510

## ; INFORMATION FOR SEQ ID NO: 128:

## SEQUENCE CHARACTERISTICS:

; LENGTH: 91 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; ORIGINAL SOURCE: peptide

; ORGANISM: Ancylostoma caninum  
US-08-486-397-128

; Query Match 83.7%; Score 36; DB 2; Length 91;

; Best Local Similarity 85.7%; Pred. No. 5.4;  
; Matches 6; Conservative 0; Mismatches

1: Indels 0; Gaps 0;

Qy 1 CEEDEYR 7

Db 59 CEEGFYR 65

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	38	88.4	638	1 GHR_HUMAN	P10912 homo sapien
2	38	88.4	638	1 GHR_MACMU	P79194 macaca mulatta
3	36	83.7	20	1 UNOS_PTHPS	P81674 pinus pinas
4	36	83.7	364	1 MYCL_HUMAN	P12524 homo sapien
5	36	83.7	368	1 MYCL_MOUSE	P10166 mus musculus
6	36	83.7	388	1 LHX9_HUMAN	Q9nq69 homo sapien
7	36	83.7	388	1 LHX9_MOUSE	Q8wuh2 mus musculus
8	36	83.7	406	1 LHX2_HUMAN	P50458 homo sapien
9	36	83.7	406	1 LHX2_MOUSE	P9z0s2 mus musculus
10	36	83.7	426	1 LHX2_RAT	P36198 ratulus norvegicus
11	35	81.4	372	1 COLB_ARATH	Q8ses5 arabidopsis thaliana
12	34	79.1	599	1 BAL_MOUSE	P04285 mus musculus
13	34	79.1	612	1 BAL_RAT	P01882 ratulus norvegicus
14	34	79.1	707	1 NRDD_HAEIN	P43752 haemophilus suis
15	33	76.7	86	1 Y576_ARCFU	Q29679 archaeoglobus fulgidus
16	32	74.4	154	1 PYRL_BUCAI	P57451 buchnera apicalis
17	32	74.4	367	1 LHX4_HUMAN	Q96992 homo sapien
18	32	74.4	367	1 LHX4_MOUSE	P53776 mus musculus
19	32	74.4	395	1 HH1R_BOVIN	P53412 gallus gallus
20	32	74.4	415	1 THD1_BAICH	P30556 bos taurus
21	32	74.4	486	1 HH1R_RAT	Q9kc63 bacillus halophilus
22	32	74.4	487	1 HH1R_HUMAN	P31390 ratulus norvegicus
23	32	74.4	488	1 HH1R_CAVPO	P35367 homo sapien
24	32	74.4	488	1 SKO1_YEAST	P31389 cavia porcellus
25	32	74.4	491	1 HH1R_BOVIN	P70174 mus musculus
26	32	74.4	581	1 NET1_CHICK	P30556 gallus gallus
27	32	74.4	606	1 PEP0_LACHE	Q90923 gallus gallus
28	32	74.4	647	1 SKO1_YEAST	Q90922 gallus gallus
29	32	74.4	647	1 NEPA1_DRONE	Q90210 lactobacillus
30	32	74.4	727	1 COPP_SCARFO	Q24567 drosoephila melanogaster
31	32	74.4	796	1 EPAA2_HUMAN	P29317 homo sapien
32	32	74.4	976	1 EPAS5_RAT	P34757 ratulus norvegicus
33	32	74.4	1005	1 EPAS5_RAT	

## ALIGNMENTS

RESULT 1					
GHR_HUMAN	ID	GHR_HUMAN	STANDARD;	PRT;	638 AA.
AC	P10912;				
DT	01-JUL-1989	(Rel. 11, Created)			
DT	01-JUL-1989	(Rel. 11, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Growth hormone Receptor precursor (GH receptor) (Serum binding protein).				
GN	GHR.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RC	TISSUE-Liver				
RX	MEDLINE=88065896; PubMed=2825030;				
RA	Leung D.W., Spencer S.A., Cachianes G., Hammonds R.G., Collins C., Hennel W.J., Barnard R., Waters M.J., Wood W.I.; "Growth hormone receptor and serum binding protein: purification, cloning and expression"; [1]				
RT	Nature 330:537-543(1987).				
RL	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90046742; PubMed=2813379;				
GA	Godowski P.J., Leung D.W., Meacham L.R., Galgani J.P., Hellmiss R., Keret R., Rotwein P.S., Parks J.S., Laron Z., Wood W.I.; "Characterization of the human growth hormone receptor gene and demonstration of a partial gene deletion in two patients with Laron-type dwarfism"; [2]				
RT	Proc. Natl. Acad. Sci. U.S.A. 86:8083-8087(1989).				
RL	[3]				
RP	DISULFIDE BONDS.				
RX	MEDLINE=90123957; PubMed=2406245;				
RA	Bourrel J.M., McKerrin M.G., Bass S., Brochier M., "Laron dwarfism in a child with a partial deletion of the extracellular domain of the growth hormone receptor gene"; [4]				
RA	Bourrel J.M., Light D.R., Wells J.A.; "The human growth hormone receptor gene deletion in two patients with Laron dwarfism"; [5]				
RT	Proc. Natl. Acad. Sci. U.S.A. 86:3111-3115(1989).				
RL	[6]				
RP	VARIANT LARON DWARFISM SER-114.				
RX	MEDLINE=89334829; PubMed=2779634;				
RA	Anslem S., Duquesnoy P., Attree O., Novelli G., Bousmina S., Postel-Vinay M.-C., Goossens M.; "Laron dwarfism and mutations of the growth hormone-receptor gene"; [7]				
RA	Valleix S., Goossens M.; "Spectrum of growth hormone receptor mutations and associated haplotypes in Laron syndrome"; [8]				
RT	Hum. Mol. Genet. 2:355-359(1993).				

VARIANT LARON DWARFISM HIS-170.  
 MEDLINE=94185615; PubMed=8137822;  
 RX Dusquesnoy P., Sobrier M.-L., Duriez B., Dasstor F., Buchanan C.R.,  
 RA Savage M.O., Preece M.A., Craescu C.T., Blouquit Y., Goossens M.,  
 RA Amselem S.;  
 RT "A single amino acid substitution in the exoplasmic domain of the  
 human growth hormone (GH) receptor confers familial GH resistance  
 (Laron syndrome) with positive GH-binding activity by abolishing  
 receptor homodimerization.";  
 RT receptor homodimerization.  
 RL EMBO J. 13:1386-1395(1994).  
 RN [7]  
 RP VARIANT IDIOPATHIC SHORT STATURE LYS-62; CYS-179 AND ASP-242.  
 RX MEDLINE=96013502; PubMed=7565946;  
 RA Godard A.D., Covello R., Luh S. M., Clackson T., Attie K.M.,  
 RA Gesundheit N., Rundle A.C., Wells J.A., Carlsson L.M.S.;  
 RT "Mutations of the growth hormone receptor in children with idiopathic  
 short stature.";  
 RL New Engl. J. Med. 333:1093-1098(1995).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 19-256.  
 RX MEDLINE=92196577; PubMed=1549797;  
 RA de Vos A.M., Uitsch M., Kossiakoff A.A.;  
 "Human growth hormone and extracellular domain of its receptor:  
 crystal structure of the complex.";  
 SCIENCE 255:306-312(1992).  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 19-256.  
 RX MEDLINE=97113023; PubMed=8949276;  
 RA Sundstrom M., Lundqvist T., Roedin J., Giell L.B., Milligan D.,  
 RA Norstedt G.;  
 RT "Crystal structure of an antagonist mutant of human growth hormone,  
 G120R, in complex with its receptor at 2.9 Å resolution.";  
 RL J. Biol. Chem. 271:32197-32203(1996).  
 CC --! SIMILARITY: THIS IS A RECEPTOR FOR PITUITARY GLAND GROWTH HORMONE.  
 CC --! SUBUNIT: HOMODIMER.  
 CC --! SUBCELLULAR LOCATION: Type I membrane protein.  
 CC --! DISEASE: DEFICIENCY IN GHR IS THE CAUSE OF PITUITARY DWARFISM II  
 CC (LARON-TYPE PITUITARY DWARFISM OR LARON SYNDROME (LS)). IT ALSO  
 CC CAUSES IDIOPATHIC SHORT STATURE.  
 CC --! SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC --! SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 EMBL; X06562; CAA29808 1;  
 EMBL; M28466; AAA5555..1; JOINED.  
 EMBL; M28458; AAA5555..1; JOINED.  
 DR EMBL; M28459; AAA5555..1; JOINED.  
 DR EMBL; M28460; AAA5555..1; JOINED.  
 DR EMBL; M28461; AAA5255..1; JOINED.  
 DR EMBL; M28462; AAA5555..1; JOINED.  
 DR EMBL; M28463; AAA5555..1; JOINED.  
 DR EMBL; M28464; AAA5255..1; JOINED.  
 DR EMBL; M28465; AAA5255..1; JOINED.  
 DR PIR; A33991; S04530.  
 DR PIR; A33991; S04530.  
 DR PDB; 3HHR; 30 APR-94.  
 DR PDB; 1HWG; 19-NOV-97.  
 DR PDB; 1HWH; 19-NOV-97.  
 DR PDB; 1AXI; 28-JAN-98.  
 DR PDB; 1A22; 29-APR-98.  
 DR Genew; HGNC:463; GHR.  
 DR MIM; 600946; -.  
 DR MIM; 262300; -.  
 DR InterPro; IPR002996; CRIA.  
 DR InterPro; IPR003961; FN\_III.  
 DR InterPro; IPR003526; Hemtopoptn\_L\_F1.

DR Pfam; PF00041; fn3; 1.  
 DR SMART; SM00050; FN; 1.  
 DR PROSITE; PS01152; HEMATopo\_REC\_L\_F1; 1.  
 KW Receptor; Transmembrane; Glycoprotein; Signal; 3D-structure;  
 KW Disease mutation.  
 SIGNAL 1 18  
 FT CHAIN 19 638  
 FT DOMAIN 19 264  
 FT TRANSMEM 265 288  
 FT DOMAIN 289 638  
 FT DOMAIN 145 66  
 FT DISULFID 101 56  
 FT DISULFID 126 140  
 FT CARBOHYD 46 46  
 FT CARBOHYD 115 115  
 FT CARBOHYD 156 156  
 FT CARBOHYD 161 161  
 FT CARBOHYD 200 200  
 FT VARIANT 62 62  
 FT VARIANT 89 89  
 FT VARIANT 114 114  
 FT VARIANT 143 143  
 FT VARIANT 162 162  
 FT VARIANT 170 170  
 FT VARIANT 179 179  
 FT VARIANT 229 229  
 FT VARIANT 242 242  
 FT VARIANT 544 544  
 FT STRAND 53 58  
 FT STRAND 64 68  
 FT STRAND 82 88  
 FT STRAND 99 100  
 FT TURN 104 107  
 FT TURN 109 110  
 FT STRAND 111 114  
 FT TURN 116 117  
 FT STRAND 124 124  
 FT TURN 132 133  
 FT STRAND 134 142  
 FT HELIX 143 145  
 FT STRAND 147 147  
 FT STRAND 153 162  
 FT TURN 164 165  
 FT STRAND 168 176  
 FT TURN 179 180  
 FT TURN 183 186  
 FT STRAND 190 198  
 FT TURN 199 200  
 FT STRAND 205 206  
 FT STRAND 210 210  
 FT STRAND 214 221  
 FT TURN 222 223  
 FT STRAND 225 234  
 FT STRAND 247 250  
 SQ SEQUENCE 638 AA; EAFT77EADE4787822 CRC64;  
 Query Match 88.4%; Score 38; DB 1; Length 638;  
 Best Local Similarity 100.0%; Pred. No. 6;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 259 CEEDFY 264

RESULT 2

GHR\_MACMU STANDARD; PRT; 638 AA.

ID GHR\_MACMU

AC P79194;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Growth hormone receptor precursor (GH receptor) (Serum binding protein).

GN Macaca mulatta (Rhesus macaque).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Cercopitheciidae; Cercopithecinae; Macaca.

NCBI\_TAXID=9544;

RN RP SEQUENCE FROM N.A. PubMed=9228076;

RX MEDLINE=97373601; PubMed=9228076;

RA Martini J.F., Pazet A., Guézennec C.Y., Edery M., Postel-Vinay M.C., Kelly P.A.:

RA "Monkey growth hormone (GH) receptor gene expression. Evidence for two mechanisms for the generation of the GH binding protein.";

RL J. Biol. Chem. 272:18951-18958(1997).

CC -- FUNCTION: THIS IS A RECEPTOR FOR PITUITARY GLAND GROWTH HORMONE.

CC -- SUBUNIT: HOMODIMER (BY SIMILARITY).

CC -- SUBCELLULAR LOCATION: TYPE I membrane protein.

CC -- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.

CC -- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and thus statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR U84589; AA847702.1; -.

DR HSSP\_P10912; TAXI.

DR InterPro; IPR00396; CRIA.

DR InterPro; IPR003961; FN\_III.

DR PROSITE; PS01352; HEMATopo\_REC\_L\_F1.

KW Receptor; Transmembrane; Glycoprotein; Signal.

FT SIGNAL 1 18 POTENTIAL.

FT CHAIN 19 638 GROWTH HORMONE RECEPTOR.

FT DOMAIN 19 264 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 285 288 POTENTIAL.

FT DOMAIN 289 638 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 145 252 FIBRONECTIN TYPE-III.

FT DISULFID 56 66 BY SIMILARITY.

FT DISULFID 101 112 BY SIMILARITY.

FT DISULFID 126 140 BY SIMILARITY.

SQ SEQUENCE 638 AA; 71327 MW; 1F8A5530 125F8E CRC64;

Query Match 88.4%; Score 38; DB 1; Length 638; Best Local Similarity 100.0%; Pred. No. 6; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEEDFY 6

Db 259 CEEDFY 264

RESULT 3

UN05\_PINPS STANDARD; PRT; 20 AA.

AC P81674;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Unknown protein from 2D-page of needles (N147) (Fragments).

OS Pinus pinaster (Maritime pine).

OC Eukaryota; Viridiplantae; Streptophytta; Embryophytta; Tracheophytta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.

OC NCBI\_TAXID=71647;

RN [1]

RP SEQUENCE.

RC TISSUE\_Needle;

RX MEDLINE=9974088; PubMed=10344291;

RA Costa P., Blonneau C., Bauw G., Dubois C., Bahrman N., Kremer A., Frigerio J.-M., Plomion C.;

RA "Separation and characterization of needle and xylem maritime pine proteins." Electrophoresis 20:1098-1108(1999).

RL MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN PROTEIN IS: 5.6, ITS MW IS: 36 kDa.

CC

FT NON\_TER 1 1

FT NON\_CONS 11 12

FT VARIANT 13 13 F -> I.

FT VARIANT 14 14 Y -> E.

FT VARIANT 15 15 R -> K.

FT NON\_TER 20 20

FT SEQUENCE 20 AA; 2438 MW; 9F4E4678E088C298 CRC64;

Query Match 83.7%; Score 36; DB 1; Length 20; Best Local Similarity 71.4%; Pred. No. 0.43; Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEEDFY 7

Db 9 CDKDFYR 15

RESULT 4

MYCL\_HUMAN STANDARD; PRT; 364 AA.

ID MYCL\_HUMAN

AC P12524; ORNLUE9;

AC P12524; ORNLUE9;

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE L-myc-1 proto-oncogene protein.

GN MYCL OR MYCL.

OS Homo sapiens (Human).

OC Metacozia; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=88094386; PubMed=2827002;

RA Kaye F., Battye J., Nau M., Brooks B., Seifert E., de Greve J., Barter M., Saussville E., Minna J.;

RA "Structure and expression of the human L-myc gene reveal a complex pattern of alternative mRNA processing." Mol. Cell. Biol. 8:186-195(1988).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=88112807; PubMed=3322939;

RA DePinho R.A., Hatton K.S., Tesiaye A., Yancopoulos G.D., Alt F.W.; "The human myc gene family: structure and activity of L-myc and an L-myc pseudogene." Genes Dev. 1:1311-1326(1987).

RN [3]

RP SEQUENCE FROM N.A.

RL Ellington A.;

RA Submitted (FEB 2000) to the EMBL/GenBank/DDBJ databases.

CC SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER BHlh protein binds DNA as an heterodimer with MAX.

CC -|- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHlh) FAMILY OF TRANSCRIPTION FACTORS. BHlh-ZIP SUBAMILY.

CC



RESULT 8							
LHX2_HUMAN		STANDARD;		PRT:			
ID	LHX2_HUMAN	P5058;	O95860;	406 AA.			
AC		01-OCT-1996 (Rel. 34, Created)					
DT	30-MAY-2000 (Rel. 39, Last sequence update)						
DT	15-JUN-2002 (Rel. 41, Last annotation update)						
DE	LIM Homeobox Protein Lhx2 (Homeobox protein Lhx-2)						
DN	LHX2 OR Lh2						
CC	-!- PRODUCED BY ALTERNATIVE SPLICING.						
CC	-!- SIMILARITY: CONTAINS 1 HOMEOBOX DOMAIN.						
CC	-!- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC IONS.						
CC	-----						
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).						
CC	-----						
DR	EMBL; AF134761; AA030110; 1;						
DR	EMBL; AF113528; AAD22008; 1;						
DR	EMBL; AJ243851; CAB59908; 1;						
DR	EMBL; AJ243852; CAB59908; 1;						
DR	EMBL; AJ243853; CAB59908; 1;						
DR	EMBL; AJ243854; CAB59908; 1;						
DR	EMBL; AJ243855; CAB59908; 1;						
DR	EMBL; AJ243856; CAB59908; 1;						
DR	EMBL; AJ243857; CAB59908; 1;						
DR	EMBL; AJ243858; CAB59908; 1;						
DR	EMBL; AJ243859; CAB59908; 1;						
DR	EMBL; AJ243854; CAB59909; 1;						
DR	EMBL; AJ243855; CAB59909; 1;						
DR	EMBL; AJ243856; CAB59909; 1;						
DR	EMBL; AJ243857; CAB59909; 1;						
DR	HSSP; P06601; 1FJL;						
DR	TRANSFAC; T04192; -;						
DR	TRANSFAC; T04195; -;						
DR	MGD; MGI:1316721; Lhx9.						
DR	InterPro; IPR0011356; Homeobox.						
DR	InterPro; IPR001781; LIM.						
DR	PFam; PF00066; homeobox.						
DR	PFam; PF00412; LIM.						
DR	ProtDom; PD000010; Homeobox.						
DR	ProDom; PD000094; LIM.						
DR	SMART; SM00389; HOX.						
DR	SMART; SM00132; LIM.						
DR	PROSITE; PS00027; HOMEOBOX_1.						
DR	PROSITE; PS50071; HOMEOBOX_2.						
DR	PROSITE; PS00078; LIM_DOMAIN_1.						
DR	PROSITE; PS50033; LIM_DOMAIN_2.						
KW	Homeobox; DNA binding; Nuclear protein; Repeat; LIM domain;						
KW	Metal-binding Zinc; Alternative splicing.						
FT	DOMAIN 62 114 LIM 1.						
FT	DOMAIN 12 177 LIM 2.						
FT	DNA_BIND 258 317 HOMEobox.						
FT	VARSPLIC 304 388 WFGQARAKPRNLQRQENGQVDKADGTSLLAPPAPADSGAL						
FT	FT TPPGATTLDLTNTVYTTVIVNMDSHEPGSQSQTLL						
FT	FT NLF -> GEQQLGHQSOTSRRLKIP (IN ISOFORM ALPHA).						
FT	CONFLICT 49 49 A -> T (IN REF. 2).						
FT	CONFLICT 153 153 S -> F (IN REF. 2).						
SQ	SEQUENCE 388 AA: 42986 MW: C2D326A8DB1B32 CRC64;						
Query Match	83 %	Score 36;	DB 1;	Length 388;			
Best Local Similarity	71.4 %	Pred. No. 8.8;					
Matches	5	Conservative 2;	Mismatches 0;	Indels 0;	Gaps 0;		
Qy	1 CEEDFYR 7						
	1:1:1:1						
Db	111 CKEDYYR 117						



Qy	1	CEEDFYR	7	Best Local Matches	5;	Similarity	71.4%;	Pred. No.	9.7;
Db	102	CKEDYR	108			Conservative	2;	Mismatches	0;
<b>RESULT 10</b>									
LHX2_RAT		STANDARD;		PRT;	426	AA.			
ID	LHX2_RAT								
AC	P36198;								
DT	01-JUN-1994	(Rel. 29, Created)							
DT	01-JUN-1994	(Rel. 29, Last sequence update)							
DT	15-JUN-2002	(Rel. 41, Last annotation update)							
DE	LIM/homeobox protein Lhx2 (Homeobox protein LH-2).								
OS	Rattus norvegicus (Rat).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Murinae; Rattus.								
OC	NCBI_TAXID=10116;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE-Brain;								
RX	MEDLINE=93126348; PubMed=2678338;								
RA	Xu Y., Baldassare M., Fisher P., Rathbun G., Oltz E.M., Yancopoulos G.D., Jessell T.M., Alt F.W.;								
RA	RT "LH-2: a LIM/homeodomain gene expressed in developing lymphocytes and neural cells.";								
RL	proc. Natl. Acad. Sci. U.S.A. 90:227-231(1993).								
CC	-!- FUNCTION: TRANSCRIPTIONAL REGULATORY PROTEIN INVOLVED IN THE CONTROL OF CELL DIFFERENTIATION IN DEVELOPING LYMPHOID AND NEURAL CELL TYPES.								
CC	-!- SUBCELLULAR LOCATION: Nuclear (Probable).								
CC	-!- TISSUE SPECIFICITY: FOUND IN DISCRETE REGIONS OF THE DEVELOPING CNS, PRIMARILY IN DINEPHRAGIC AND TELENCEPHALIC STRUCTURES AND A SUBSET OF LYMPHOID TISSUES. ALSO FOUND IN EMBRYONIC SPINAL CHORD AND FETAL LIVER.								
CC	-!- DEVELOPMENTAL STAGE: EXPRESSED IN DEVELOPING LYMPHOCYTES AND NEURAL CELLS. MAXIMAL EXPRESSION IS FOUND IN PRE-B-LYMPHOCYTES.								
CC	-!- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN. THE LIM DOMAIN BINDS 2 ZINC IONS.								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).								
CC	DR EMBL: L06804; :- NOT_ANNOTATED_CDS.								
CC	DR HSSP: P06601; IFUJ.								
CC	DR InterPro: IPRO01356; Homeobox.								
CC	DR InterPro: IPR01781; LIM.								
CC	DR PFAM: PF00046; homeobox; 1.								
CC	DR ProDom: PD000010; Homeobox; 1.								
CC	DR ProDom: PD000094; LIM; 2.								
CC	SMART: SM00389; HOX; 1.								
CC	SMART: SM00132; LIM; 2.								
CC	DR PROSITE: PS00027; HOMEobox_1; 1.								
CC	DR PROSITE: PS00476; LIM_DOMAIN_1; 2.								
CC	DR PROSITE: PS50023; LIM_DOMAIN_2; 2.								
CC	DR PROSITE: PS50071; HOMEobox_2; 1.								
CC	KW Homeobox: DNA-binding; Nuclear protein; Repeat; LIM domain; Metal I-binding; Zinc; transcription regulation.								
CC	FT DOMAIN 52 104 LIM 1.								
CC	FT DOMAIN 114 167 LIM 2.								
CC	FT DOMAIN 264 323 HOMEobox.								
CC	FT DOMAIN 305 321 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).								
CC	SEQUENCE 426 AA; 47418 MW; DC8PA3DBA572BB40 CRC64;								
Query Match	83.7%	Score 36;	DB 1;	Length 426;					

DR	PROSITE; PS50119; ZF_BBOX; 2.	Repeat; Multigene family.	DR	Pfam; PF00135; Coesterase; 1.
KW	Zinc-finger; Nuclear protein; 47	B BOX-TYPE 1.	DR	PROSITE; PS01122; CARBOXYLESTERASE_B_1; 1.
FT	ZN_FING 5	B BOX-TYPE 2 (ATOPICAL).	DR	PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
FT	ZN_FING 48	POLY-ASN.	KW	Hydrolase; Serine esterase; Lipid degradation; Glycoprotein;
FT	77	POLY-SER.	Repeat; Signal.	
FT	83		FT	SIGNAL 1 20 BY SIMILARITY.
FT	90		FT	CHAIN 21 599 BY BILE-SALT-ACTIVATED LIPASE.
SEQUENCE	84	MW: 188F1BEB283D7479 CRC64;	FT	ACT_SITE 214 214 BY SIMILARITY.
SEQUENCE	372 AA;	40754 MW;	FT	ACT_SITE 340 340 BY SIMILARITY.
Query Match	81.4%	Score 35; DB 1; Length 372;	FT	ACT_SITE 455 455 BY SIMILARITY.
Best Local Similarity	83.3%	Pred. No. 13;	FT	DISULFID 84 100 BY SIMILARITY.
Matches 5;	conservative 1;	Mismatches 0;	FT	DISULFID 266 277 BY SIMILARITY.
QY	1 CEDFY 6	Gaps 0;	FT	DOMAIN 559 588 4 X 11 AA TANDEM REPEATS, O-GLYCOSYLATED REGION.
Db	181 CEDFY 186		FT	REPEAT 559 569 1.
			FT	REPEAT 570 580 2.
			FT	REPEAT 581 588 3.
RESULT 12	BAL_MOUSE	STANDARD;	FT	N-LINKED (GLCNAC. . .) (POTENTIAL).
ID	BAL_MOUSE	PRT;	FT	CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
AC	Q64385;	PRT;	FT	CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (POTENTIAL).
DE	01-NOV-1997 (Rel. 35, Created)	599 AA.	SQ	SEQUENCE 599 AA; 65813 MW; 9E4428FDFA8602E CRC64;
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DE	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	Bile-salt-activated lipase precursor (EC 3.1.1.3) (BAL)			
DE	(Bile-salt-stimulated lipase) (BSSL) (Carboxy ester lipase) (Sterol esterase) (Cholesterol esterase) (Pancreatic lysophospholipase).			
DE	CEL OR LIP1.			
GN	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			RESULT 13
RP	SEQUENCE FROM N.A.		ID	BAL_RAT STANDARD; PRT; 612 AA.
RC	STRAIN=BALB/c; TISSUE=Mammary gland; MEDLINE=16096331; PubMed=8522186;		AC	P07882; P14722;
RX	Mackay K, Law R, M;		DT	01-AUG-1988 (Rel. 08, Created)
RA	"Characterization of the mouse pancreatic/mammary gland cholesterol-encoding cDNA and gene." RT		DT	01-APR-1990 (Rel. 14, Last sequence update)
RT	Lidner A.-S., Kannius M., Lundberg L., Bjursell G., Nilsson J.; RT		DT	01-OCT-1996 (Rel. 34, Last annotation update)
RL	"Molecular cloning and characterization of the mouse carboxyl ester lipase gene and evidence for expression in the lactating mammary gland." RT		DE	Bile-salt-activated lipase precursor (EC 3.1.1.3) (BAL)
RN	Genomics 29:115-122(1995).		DE	(Bile-salt-stimulated lipase) (BSSL) (Carboxy ester lipase) (Sterol esterase) (Cholesterol esterase) (Pancreatic lysophospholipase).
RP	SEQUENCE FROM N.A.		GN	CE.
RC	STRAIN=BALB/c; TISSUE=Lactating mammary gland; MEDLINE=16079098; PubMed=8530060;		OS	Rattus norvegicus (Rat).
RX	RA		OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RA	Lidner A.-S., Kannius M., Lundberg L., Bjursell G., Nilsson J.; RT		OX	NCBI_TaxID=10116;
RT	"Molecular cloning and characterization of the mouse carboxyl ester lipase gene and evidence for expression in the lactating mammary gland." RT		RN	[1]
RL	Genomics 29:115-122(1995).		RP	SEQUENCE FROM N.A.
CC	-!- FUNCTION: CATALYZES FAT AND VITAMIN ABSORPTION. ACTS IN CONCERT WITH PANCREATIC LIPASE AND COLIPASE FOR THE COMPLETE DIGESTION OF DIETARY TRIGLYCERIDES (BY SIMILARITY).		RC	SEQUENCE FROM N.A.
CC	-!- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a fatty acid anion.		RC	SEQUENCE FROM N.A.
CC	-!- CATALYTIC ACTIVITY: A sterol ester + H(2)O = a sterol + a fatty acid.		RA	Fontaine R.N., Stratowa C., Rutter W.J.;
CC	-!- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.		RT	Isolation of full-length putative rat lysophospholipase cDNA using RT
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).		RT	improved methods for mRNA isolation and cDNA cloning.";
CC	DR	EMBL; U33169; AAA92088; 1;	RA	RA
DR	HSSP; U37386; ANC52279; 1;		RT	Dierscio L.P., Fontaine R.N., Hui D.Y.;
DR	P30122; BCCE.		RA	"Identification of the active site serine in pancreatic cholesterolemia esterase by chemical modification and site-specific mutagenesis.";
DR	MGI; MGI-88374; Cel.		RT	J. Biol. Chem. 265:16801-16806(1990).
DR	InterPro; IPR002018; Carboxylesterases.		RL	
DR	InterPro; IPR000379; Ser_enzrs_site.		RN	
DR	ACTIVE SITE SER-214.		RN	
			RP	Medline=91005095; Pubmed=2211595;
			RX	Dierscio L.P., Fontaine R.N., Hui D.Y.;
			RA	"Identification of the active site serine in pancreatic cholesterolemia esterase by chemical modification and site-specific mutagenesis.";
			RT	J. Biol. Chem. 265:16801-16806(1990).

RN [5]	ACTIVE SITE HIS-455.	Db 361 EEDFYR 366
RX MEDLINE=91154187; PubMed=1999399;		
RA Dipersio L.P., Fontaine R.N., Hui D.Y.;		
RT Site specific mutagenesis of an essential histidine residue in pancreatic cholesterol esterase.;"		
RT Site		
RL J. Biol. Chem. 266:4033-4036(1991)		
CC -!- FUNCTION: CATALYZES FAT AND VITAMIN ABSORPTION. ACTS IN CONCERT WITH PANCREATIC LIPOASE AND COLIPASE FOR THE COMPLETE DIGESTION OF DIETARY TRIGLYCERIDES.		RESULT 14 NRDD_HAEIN ID NRDD_HAEIN AC F43752; DT 01-NOV-1995 (Rel. 32, Created) DT 01-NOV-1995 (Rel. 32, Last sequence update) DT 15-JUN-2002 (Rel. 41, Last annotation update) DE Anaerobic ribonucleoside-triphosphate reductase (EC 1.17.4.2). GN NRDD OR HGIC075. OS Haemophilus influenzae
CC -!- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a fatty acid anion.		
CC -!- CATALYTIC ACTIVITY: A sterol ester + H(2)O = a sterol + a fatty acid.		
CC -!- ENZYME REGULATION: ACTIVATED BY BILE SALTS CONTAINING A 7-HYDROXYL GROUP.		
CC -!- TISSUE SPECIFICITY: SYNTHESIZED PRIMARILY IN THE PANCREAS AND THEN TRANSPORTED TO THE INTESTINE.		
CC -!- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.		
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).		
CC DR X16054; CAA34189.1 ;		
DR EMBL: M15853; AAA41540.1 ;		
DR EMBL: M69157; AAB46376.1 ;		
DR PIR: A34967; A34967.		
DR PIR: A26603; A26603.		
DR HSSP: P30122; 2BCE.		
DR InterPro: IPR002018; CarboxylesteraseB		
DR InterPro: IPR003079; Ser.estrs_site.		
DR Pfam: PF00135; Coesterase_1.		
DR PROSITE: PS00122; CARBOXYLESTERASE_B-1; 1.		
DR PROSITE: PS00122; CARBOXYLESTERASE_B-2; 1.		
KW Hydrolase; Serine esterase; Lipid degradation; Glycoprotein;		
KW Repeat; Signal	1 20	
FT SIGNAL	1 20	
FT CHAIN	21 612	BILE-SALT-ACTIVATED LIPASE.
FT ACT_SITE	214 214	
FT ACT_SITE	340 340	BY SIMILARITY.
FT ACT_SITE	455 455	BY SIMILARITY.
FT DISULFID	84 100	BY SIMILARITY.
FT DISULFID	266 277	
FT CARBOHYD	207 207	N-LINED (GLCNAC. . .) (POTENTIAL).
FT DOMAIN	556 599	4 X 11 AA TANDEM REPEATS, O-GLYCOSYLATED REGION.
FT REPEAT	556 566	1.
FT REPEAT	567 577	
FT REPEAT	578 588	2.
FT REPEAT	589 599	3.
FT MUTAGEN	440 440	H->Q: NO EFFECT ON ACTIVITY.
FT MUTAGEN	455 455	H->Q, R,A,S,D: ABOLISH ACTIVITY.
FT CONFLICT	26 26	V -> L (IN REF. 2).
FT CONFLICT	154 154	G -> A (IN REF. 2).
FT CONFLICT	217 217	A -> G (IN REF. 2).
FT CONFLICT	219 219	S -> I (IN REF. 2).
FT CONFLICT	419 419	M -> T (IN REF. 3).
FT CONFLICT	513 513	T -> M (IN REF. 2 AND 3).
FT CONFLICT	576 577	GG -> VV (IN REF. 3).
FT CONFLICT	608 609	GP -> VA (IN REF. 3).
FT CONFLICT	611 611	G -> A (IN REF. 3).
SQ SEQUENCE	67040 MW;	1565CE4EA71EDD2A CRC64;
Query Match	79.1%	Score 34; DB 1; Length 612;
Best Local Similarity	100.0%	Pred. No. 34;
Matches	0;	Mismatches 0; Indels 0; Gaps 0;
Qy 2 EEDFYR 7		RESULT 15 Y576_ARCFU ID Y576_ARCFU AC 029679;
Db 425 EEDFYR 430		

DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein AF0576.  
 GN AF0576.  
 OS Archaeoglobus fulgidus.  
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
 OC Archaeoglobaceae; Archaeoglobus.  
 OX NCBL\_TaxID=2234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
 RX MEDLINE=98049343; PubMed=9389475;  
 RA Klein H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,  
 RA Richardson D.L., Kerlavage A.R., Graham D.E., Myrordes N.C.,  
 RA Fleischmann R.D., Quackenbush J., Sutton G.G., Gill S.,  
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
 RA Overbeek R., Gocayne J.D., Weidman J.P., McDonal L., Utterback T.,  
 RA Cotton M.D., Spriggs T., Ariyach P., Kaine B.P., Styles S.M.,  
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
 RA Venter J.C.;  
 RT "The complete genome sequence of the hyperthermophilic, sulphate-  
 reducing archaeon *Archaeoglobus fulgidus*."  
 RL Nature 390:364-370(1997).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC use by non profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL: AE00105; AAB90675.1; -.  
 DR TIGR; AF0576; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 86 AA; 10018 MW; 03AA3D19DAB8000C CRC64;  
 -----

Query Match 76.7%; Score 33; DB 1; Length 86;

Best Local Similarity 71.4%; Pred. No. 7.1; Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CEEFYR 7  
 |||:  
 Db 67 CEEFYR 73

Search completed: February 12, 2003, 11:44:59

time : 7.53333 secs

Result No.	Score	Query	Match	Length	DB	ID	Description
1	38	88 4	153	16	Q8R8G6		08zg6 salmonella
2	38	88 4	632	6	Q95ML5		Q8zg7 salmonella
3	38	88 4	638	6	Q9XSZ1		Q26637 strongyloce
4	37	86 0	399	11	Q9QY66		Q26638 paracentrot
5	37	86 0	399	11	Q8VCQ7		Q26639 strongyloce
6	36	83 7	76	5	Q9U9B1		Q9sk1 arabidopsis
7	36	83 7	91	5	Q16938		Q97d4 clostridium
8	36	83 7	125	13	Q98SF6		Q8r45 thermoaner
9	36	83 7	206	4	Q14897		Q8yn6 anabaena sp
10	36	83 7	217	13	Q98SP7		Q9w4q9 drosophila
11	36	83 7	297	11	Q9CSG0		Q95X85 caenorhabdi
12	36	83 7	378	13	Q908B1		Q98PN8 mycoplasma
13	36	83 7	400	13	Q42108		Q94568 melanoplus
14	36	83 7	726	10	Q9M9V2		Q96JP7 homo sapien
15	35	81 4	202	16	Q9PNX2		Q95X85 caenorhabdi
16	35	81 4	219	10	Q8S056		Q98WTn7 cotesia kar

DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Growth hormone receptor.

OS *Saimiri boliviensis* (Bolivian squirrel monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.

NCBI\_TaxID=27679;  
 RN [1]

RP SEQUENCE FROM N.A. PubMed=11371582;

RA Liu J.C., Makova K.D., Adkins R.M., Gibson S., Li W.H.;  
 RT "Episodic Evolution of Human Growth Hormone Receptor.";  
 RT Species Specificity of Human Growth Hormone Receptor.;"  
 RL Mol. Biol. Evol. 18:945-953 (2001).

DR AF339051; AAK62288.1; -.

DR InterPro: IPR002986; CRLA.

DR InterPro: IPR003961; FN\_III.

DR InterPro: IPR003528; Hemtopoptn\_L\_F1.

PFAM: PF00041; fn3; 1.

SMART: SM00060; fn3; 1.

DR Receptor.

PROSITE: PS01352; HEMATOPO\_REC\_L\_F1; UNKNOWN\_1.

SEQUENCE 632 AA; 70883 MW; 440E17AF627EDA3 CRC64;

Query Match 88.4%; Score 38; DB 6; Length 632;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEEDEFY 6  
 Dd 259 CEEDEFY 264

RESULT 3  
 Q9ASZ1 PRELIMINARY; PRT; 638 AA.  
 ID Q9ASZ1; PRELIMINARY; PRT; 638 AA.  
 AC Q9ASZ1; PRELIMINARY; PRT; 638 AA.  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-MAR-2002 (TREMBLrel. 12, Last sequence update)  
 DE Growth hormone receptor.

OS *Papio anubis* (Olive baboon).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Papio.

NCBI\_TaxID=9555;  
 RN [1]

RP SEQUENCE FROM N.A. PubMed=10425448;

RA Zegopoulos G., Nathanielsz P., Hendy G.N., Goodyer C.G.;

RT "The baboon: a model for the study of primate growth hormone receptor gene expression during development.";  
 RL J. Mol. Endocrinol. 23:67-75(1999).  
 DR HSSP; P10912; IAXI.

DR InterPro: IPR002986; CRLA.

DR InterPro: IPR00361; FN\_III.

DR InterPro: IPR00328; Hemtopoptn\_L\_F1.

PFAM: PF00041; fn3; 1.

SMART: SM00060; FN3; 1.

PROSITE: PS01352; HEMATOPO\_REC\_L\_F1; UNKNOWN\_1.

KW Receptor.

SEQ 638 AA; 71407 MW; 9E250C8E303E420 CRC64;

Query Match 88.4%; Score 38; DB 6; Length 638;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEEDEFY 6  
 Dd 259 CEEDEFY 264

RESULT 4  
 Q9QY66 PRELIMINARY; PRT; 399 AA.  
 ID Q9QY66; PRELIMINARY; PRT; 399 AA.  
 AC Q9QY66; PRELIMINARY; PRT; 399 AA.  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE C1orf5.

GN ORF5.

OS *Mus musculus* (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI\_TaxID=10090;  
 RN [1]

RP SEQUENCE FROM N.A. MEDLINE=20065078; PubMed=10602999;

RA Lemmens T.H., Farnebo F., Piehl F., Merregaert J., Van de Ven W.J.M., Larsson C., Kas K.;  
 RA "Molecular characterization of human and murine cliforf5, a new member of the FAUNA gene cluster.";  
 RT RL Mamm. Genome 11:78-80(2000).  
 DR EMBL; AF119408; AAF235021; -.

DR MGII:135481; ORF6.

SEQUENCE 399 AA; 43038 MW; 6BED852632747B54 CRC64;

Query Match 86.0%; Score 37; DB 11; Length 399;  
 Best Local Similarity 85.7%; Pred. No. 21;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CEEDEFYR 7  
 Db 41 CAEDFYR 47

RESULT 5  
 Q8VCQ7 PRELIMINARY; PRT; 399 AA.  
 ID Q8VCQ7; PRELIMINARY; PRT; 399 AA.  
 AC Q8VCQ7; PRELIMINARY; PRT; 399 AA.  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DE Open reading frame 6.

OS *Mus musculus* (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI\_TaxID=10090;  
 RN [1]

RP SEQUENCE FROM N.A. RC TISSUE-LIVER;  
 RA Strausberg R.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC019440; AAH19440.1; -.

SEQUENCE 399 AA; 42995 MW; B5CA88342B56932B CRC64;

Query Match 86.0%; Score 37; DB 11; Length 399;  
 Best Local Similarity 85.7%; Pred. No. 21;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CEEDEFYR 7  
 Db 41 CAEDFYR 47

RESULT 6  
 Q9U9B1 PRELIMINARY; PRT; 76 AA.  
 ID Q9U9B1; PRELIMINARY; PRT; 76 AA.  
 AC Q9U9B1; PRELIMINARY; PRT; 76 AA.  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE Ascaris type serine protease inhibitor (Fragment).

OS *Ankylostoma ceylanicum*.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
 OC Ancylostomatoidea; Ancylostomatidae; Ancylostoma; Ancylostoma.  
 OX NCBI\_TAXID=53326;

RN [1] SEQUENCE FROM N.A.

RA Harrison L.M., Cappello M.;  
 RT "The molecular cloning of an ascaris type serine protease inhibitor  
 from adult Ancylostoma ceylanicum hookworms.";  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF172653; AAD61336.1; - .  
 DR HSSP: P56682; ICCV .

DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR002919; TIL\_Cysrich.

DR Pfam: PF01826; TIL; 1.  
 DR PROSITE: PS01186; EGF\_2; UNKNOWN\_1.

KW Protease.

FT NON-TER 1 1  
 SEQUENCE 76 AA; 8385 MW; D35FCCEF7C2088A53 CRC64;

Query Match 83.7%; Score 36; DB 5; Length 76;  
 Best Local Similarity 85.7%; Pred. No. 6.1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CEEDFYR 7  
 1 1 1 1  
 Db 44 CEEGFIR 50

RESULT 7  
 Q16938 ID Q16938 PRELIMINARY; PRT; 91 AA.

AC Q16938; DT 01-NOV-1996 (TREMBLrel. 01; Created)  
 DT 01-NOV-1996 (TREMBLrel. 01; Last sequence update)

DE Anti-coagulant protein C2 precursor (Fragment).  
 OS Ancylostoma caninum (Dog hookworm).

OC Eukaryota; Metazoa; Chromadorea; Rhabditida; Strongylida;  
 Ancylostomatoidea; Ancylostomatidae; Ancylostoma; Ancylostoma.

OX NCBI\_TAXID=29170;

RN [1] SEQUENCE FROM N.A.

RP MEDLINE=98298519; PubMed=9634780;

RA Jespers L.S., Messens J.H., De Keyser A., Beckhout D.,  
 van den Brande I., Ganssema Y.G., Lauwereys M.J., Vlaesuk G.P.,  
 Stanssens P.E.;

RA "Surface expression and ligand-based selection of cDNAs fused to  
 filamentous phage gene VI";  
 RL Biotechnology 13:378-382(1995).  
 RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=96312555; PubMed=0700900;

RA Stanssens P., Bergum P.W., Ganssema Y., Jaspers L., Laroche Y.,  
 RA Huang S., Maki S., Messens J.H., Lauwereys M., Cappello M., Hotez P.J.,  
 RA Lasters I., Vlaesuk G.P.;

RA "Anticoagulant repertoire of the hookworm Ancylostoma caninum.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:2449-2454 (1996).

DR EMBL: U30791; AAC47080.1; - .  
 DR HSSP: P56682; ICCV .

DR InterPro: IPR002919; TIL\_Cysrich.

DR Pfam: PF01826; TIL; 1.  
 DR PROSITE: PS01186; EGF\_2; UNKNOWN\_1.

KW Signal.

FT NON-TER 1 1  
 FT SIGNAL <1 7  
 FT CHAIN 8 91 POTENTIAL.  
 SEQUENCE 91 AA; 10358 MW; ANTI-COAGULANT PROTEIN C2.

Query Match 83.7%; Score 36; DB 5; Length 91;  
 Best Local Similarity 85.7%; Pred. No. 7.3; Mismatches 1; Indels 0; Gaps 0;

OC Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopus.

OX NCBI\_TAXID=8355;

RN [1] SEQUENCE FROM N.A.

RC TISSUE-HEAD;

RA Bachy I., Vernier P., Retaux S.;

RL The LIM-homeodomain family in the developing xenopus brain:  
 conservation and divergences with the mouse related to the evolution  
 of the forebrain.";

RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

CC "-! SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC  
 IONS.

DR EMBL: AJ311712; CAC35215.1; - .

DR HSSP: P329655; ICML.

DR InterPro: IPR001781; LIM.

DR PF004112; LIM; 2.

DR PRODOM: PD000094; LIM; 1.

DR SMART: SM00132; LIM; 2.

DR PROSITE: PS00478; LIM\_DOMAIN\_1; 1.

DR PROSITE: PS50023; LIM\_DOMAIN\_2; 2.

KW LIM domain; Metal-binding; Zinc.

FT NON-TER 1 1

FT NON-TER 125 125

SQ SEQUENCE 125 AA; 14283 MW; 375E42A29104D364 CRG64;

Query Match 83.7%; Score 36; DB 13; Length 125;

Best Local Similarity 71.4%; Pred. No. 10; Mismatches 2; Indels 0; Gaps 0;

CC 5; Conservative

Oy 1 CEEDFYR 7  
 1 1 1 1  
 Db 28 CKEDYR 34

RESULT 9  
 Q14897 ID Q14897 PRELIMINARY; PRT; 206 AA.

AC Q14897; DT 01-NOV-1996 (TREMBLrel. 01; Created)

DT 01-NOV-1996 (TREMBLrel. 01; Last sequence update)  
 DR L-myc protein (Similar to lung carcinoma myc related oncogene 1).

GN L-MYC.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TAXID=9606;

RN [1] SEQUENCE FROM N.A.

RC TISSUE-PLACENTA;

RX MEDLINE=8804386; PubMed=2827002;

RA Raye F., Battye J., Nau M., Brooks B., Seifert E., De Greve J.,  
 Birrer M., Sausville E., Minna J.;

RT "Structure and expression of the human L-myc gene reveal a complex  
 pattern of alternative mRNA processing.", Mol. Cell. Biol. 8:186-195(1988).

RL

RN	[2]	SEQUENCE FROM N.A.	PRT;	297 AA.
RC	TISSUE=PLACENTA;			
RA	Straubberg R.;			
RL	Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.			
DR	EMBL; M19720; AAAS59818; AAH18641; -;			
DR	EMBL; BC011864; AAH18641; -;			
DR	TRANSFAC; T02306; -;			
DR	InterPro; IPR002418; -;			
DR	Pfam; PF01056; MYC_N term; 1;			
DR	SEQUENCE 206 AA; 21766 MW;	585c9CD6c9A8EC71 CRC64;		
Query Match	Best Local Similarity 83.7%; Score 36; DB 4; Length 206;			
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;				
Qy	1 CEEDEFYR 7	PRELIMINARY;	PRT;	217 AA.
Db	1 I I I I	01-JUN-2001 (TREMBLrel. 17, Created)		
	15 CGEDFYR 21	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
		01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
RESULT 10	Xenopus laevis (African clawed frog)			
098SF7	SEQUENCE FROM N.A.			
AC	098SF7;	PRELIMINARY;	PRT;	217 AA.
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DE	Lhx9 protein (Fragment).			
GN	LHX9.			
OS	Xenopus laevis (African clawed frog)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopidae; Xenopus.			
OX	NCBI_TAXID=8355;			
RN	[1]	SEQUENCE=HEAD;		
RC	Bachy I., Vernier P., Retaux S.;			
RA	"The LIM-homeobox family in the developing xenopus brain: conservation and divergences with the mouse related to the evolution of the forebrain.";			
RT	Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.			
CC	-! SIMILARITY: CONTAINS 1 LIM DOMAIN. THE LIM DOMAIN BINDS 2 ZINC IONS.			
CC	EMBL; AJ311711; CAC35214.1; -;			
DR	InterPro; IPR001356; Homeobox.			
DR	InterPro; IPR001781; LIM.			
DR	Pfam; PF00046; homeobox; 1.			
DR	Pfam; PF00412; LIM; 2.			
DR	ProDom; PD000010; Homeobox; 1.			
DR	ProDom; PD00094; LIM; 2.			
DR	SMART; SM00389; HOX; 1.			
DR	SMART; SM00132; LIM; 1.			
DR	PROSITE; PS50071; HOMEBOX_2; 1.			
DR	PROSITE; PS500478; LIM_DOMAIN_1; 1.			
DR	PROSITE; PS50023; LIM_DOMAIN_2; 2.			
KW	LIM domain; Metal-binding; Zinc.			
FT	NON_TER 1			
FT	SEQUENCE 217 AA; 24716 MW;	6CBEB8BB595ECB851 CRC64;		
Query Match	Best Local Similarity 83.7%; Score 36; DB 13; Length 217;			
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;				
Qy	1 CEEDEFYR 7	PRELIMINARY;	PRT;	378 AA.
Db	1 : I : I : I	01-NOV-1996 (TREMBLrel. 01, Created)		
	27 CKEDFYR 33	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
RESULT 11	Homeobox protein.			
Q9CSG0				
ID	Q9CSG0;	PRELIMINARY;	PRT;	297 AA.
AC	Q9CSG0;			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	LIM homeobox protein 9 (Fragment).			
GN	Mus musculus (Mouse).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OC	NCBI_TAXID=10090;			
RN	[1]	SEQUENCE FROM N.A.		
RP	STRAIN=C57BL/6J; TISSUE=EMBRYO;			
RC	Medline=21085660; PubMed=11217831;			
RX	Kawai J., Shinagawa A., Shioda K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa M., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamamoto I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nakaido I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Garibaldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarollo J., Nobrega P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakanoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;			
RT	"Functional annotation of a full-length mouse cDNA collection.";			
RL	Nature 409:685-690 (2000);			
CC	-! SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC IONS.			
CC	EMBL; AK012930; BAB28555.1; -.			
DR	HSSP; P32965; 1CTC.			
DR	MGI; MGI:1316721; Lhx9.			
DR	InterPro; IPR001356; Homeobox.			
DR	InterPro; IPR001781; LIM.			
DR	Pfam; PF00046; homeobox; 1.			
DR	Pfam; PF00412; LIM; 2.			
DR	ProDom; PD000010; Homeobox; 1.			
DR	ProDom; PD00094; LIM; 2.			
DR	SMART; SM00389; HOX; 1.			
DR	SMART; SM00132; LIM; 1.			
DR	PROSITE; PS50071; HOMEBOX_2; 1.			
DR	PROSITE; PS500478; LIM_DOMAIN_1; 1.			
DR	PROSITE; PS50023; LIM_DOMAIN_2; 2.			
KW	LIM domain; Metal-binding; Zinc.			
FT	NON_TER 1			
FT	SEQUENCE 297 AA; 33502 MW;	3840FE3B819BB053 CRC64;		
Query Match	Best Local Similarity 83.7%; Score 36; DB 11; Length 297;			
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;				
Qy	1 CEEDEFYR 7	PRELIMINARY;	PRT;	378 AA.
Db	1 : I : I : I	01-NOV-1996 (TREMBLrel. 01, Created)		
	27 CKEDFYR 33	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
RESULT 12	Homeobox protein.			
Q90881				
ID	Q90881;	PRELIMINARY;	PRT;	378 AA.
AC	Q90881;			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DE	LIM homeobox protein 9 (Fragment).			
GN	LHX-2.			

OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RA Tremml G.; Jessell T.M.;  
 RT "Differentiation of dorsal commissural neurons defined by expression  
 of the LIM homeobox gene Lh-2: suppression by notochord grafts and  
 maintained after notochord removal";  
 RT Submitted "(OCT-1994) to the EMBL/GenBank/DDBJ databases."  
 RL -1- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC  
 CC IONS.  
 DR EMBL; L35566; AAA50258..1; -.  
 DR HSSP; P32965; 1CTL.  
 DR InterPro; IPR01556; Homeobox.  
 DR InterPro; IPR00047; RTH\_repressor.  
 DR InterPro; IPR01781; LIM.  
 DR Pfam; PF00046; homeobox; 1.  
 DR Pfam; PF00412; LIM; 2.  
 DR PRINTS; PR00031; HTREPRESSR.  
 DR PRODOM; PD000010; Homeobox; 1.  
 DR SMART; SM00094; LIM; 2.  
 DR SMART; SM00389; HOX; 1.  
 DR SMART; SM00132; LIM; 2.  
 DR PROSITE; PS00027; HOMEOBOX\_1; 1.  
 DR PROSITE; PS00071; HOMEOBOX\_2; 1.  
 DR PROSITE; PS00478; LIM\_DOMAIN\_1; 2.  
 DR PROSITE; PS0023; LIM\_DOMAIN\_2; 2.  
 KW DNA-binding; Homeobox; LIM domain; Metal-binding; Nuclear protein;  
 KW Zinc.  
 SQ SEQUENCE 378 AA; 42007 MW; 34220850FCEB2FFC CRC64;  
 Query Match 83.7%; Score 36; DB 13; Length 378;  
 Best Local Similarity 71.4%; Pred. No. 32;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 RN [1]  
 RP CEEDEFYR 7  
 DR :||:||| 101 CKEDYVR 107

---

RESULT 13  
 ID Q42108 PRELIMINARY; PRT; 400 AA.  
 AC 042108;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE LIM homeodomain.  
 GN LH-2A.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-LIMB BUD;  
 RX MEDLINE-97446014; PubMed=9299541;  
 RA Nohno T.; Kawakami Y.; Wada N.; Ishikawa T.; Ohuchi H.; Noji S.;  
 RT "Differential expression of the two closely related LIM-class homeobox  
 genes LH-2A and LH-2B during limb development.".  
 RL Biochem. Biophys. Res. Commun. 238:506-511(1997).  
 -1- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC  
 CC IONS.  
 DR EMBL; AB005882; BAA21846..1; -.  
 DR HSSP; P06601; 1FL1.  
 DR InterPro; IPR01356; Homeobox.

---

RESULT 14  
 ID Q9M9Y2 PRELIMINARY; PRT; 726 AA.  
 AC Q9M9Y2;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE F4H5..17 protein.  
 GN F4H5..17.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC euros II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TAXID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fedderspiel N.A.; Palm C.J.; Conway A.B.; Conn L.; Hansen N.F.;  
 RA Araujo R.; Huizar L.; Rowley D.; Buehler E.; Dunn P.;  
 RA Gonzalez A.; Kremenetskaia I.; Kim C.; Lenz C.; Li J.; Liu S.;  
 RA Luers S.; Schwartz J.; Shinn P.; Toriumi M.; Vysotskaya V.S.;  
 RA Walker M.; Yu G.; Ecker J.; Theologis A.; Davis R.W.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AC011001; AAF63144..1; -.  
 DR InterPro; IPR004332; MuDR.  
 DR InterPro; IPR004862; MurA.  
 DR InterPro; IPR01878; zf\_fCCHC.  
 DR Pfam; PF03108; MuDR; 1.  
 DR Pfam; PF03163; MURA; 1.  
 DR SMART; SM00398; zf-CCCHC; 1.  
 DR SMART; SM00343; Znf\_CCHC; 1.  
 SQ SEQUENCE 726 AA; 82359 MW; 6B9A6133C3BBF5B2 CRC64;  
 Query Match 83.7%; Score 36; DB 10; Length 726;  
 Best Local Similarity 85.7%; Pred. No. 62;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

---

RESULT 15  
 ID Q9PNX2 PRELIMINARY; PRT; 202 AA.  
 AC Q9PNX2;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Hypothetical protein Cj0963.

GN CJ0963  
OS Campylobacter jejuni.  
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
CC Campylobacter  
OX NCBI\_TaxID=19;;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=20150912; PubMed=10688204;  
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,  
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holroyd S.,  
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,  
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,  
RA Whitehead S., Barrell B.G.;  
RT "The genome sequence of the food-borne pathogen *Campylobacter jejuni*  
reveals hypervariable sequences.";  
RL Nature 403:665-668 (2000).  
EMBL: ALI39076; CAB73220.1.  
DR Hypothetical protein; Complete proteome.  
KW SEQUENCE 202 AA; 23926 MW;  
SQ

Query Match 81.48; Score 35; DB 16; Length 202;  
Best Local Similarity 83.3%; Pred. No. 26;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CEEDEFY 6  
1:1111  
Db 99 CDEDEFY 104

Search completed: February 12, 2003, 11:47:00  
Job time : 54.7333 secs